Economic Review

Federal Reserve Bank of San Francisco

Summer 1990 Number 3

Table of Contents

The Federal Reserve Bank of San Francisco's Economic Review is published quarterly by the Bank's Research Department under the supervision of Jack H. Beebe, Senior Vice President and Director of Research. The publication is edited by Barbara A. Bennett. Design, production, and distribution are handled by the Public Information Department, with the assistance of Karen Rusk and William Rosenthal.

For free copies of this and other Federal Reserve publicatons, write or phone the Public Information Department, Federal Reserve Bank of San Francisco, P.O. Box 7702, San Francisco, California 94120. Phone (415) 974-2163.

Do Interest Rates Still Affect Housing?

Randall J. Pozdena

Vice President, Federal Reserve Bank of San Francisco.

The author wishes to thank Rachel Long and Deborah Martin for their skilled and patient research assistance. Editorial committee members were Carolyn Sherwood-Call, Jonathan Neuberger, and Adrian Throop.

Activity in the housing market traditionally has been very sensitive to changes in interest rates. This sensitivity has had important implications both for participants in the housing industry and monetary policy. Theory suggests, however, that financial innovation and deregulation in recent years may have altered the link between housing and interest rates. In this paper, the theoretical linkages are discussed and studied empirically for the periods before and after 1983. A significant difference in the strength and nature of the linkages is revealed.

The housing market historically has displayed pronounced cycles in investment activity. In the last 30 years, for example, the variation in the volume of new housing starts has been 1½ times greater than the variation in GNP over the same period.[1] Economists and central bankers long have been interested in understanding these fluctuations for several reasons. First, the cycles in housing activity empirically have been useful leading indicators of the general business cycle. In most recessions, residential fixed investment appears to have led both declines in business investment and GNP.[2] Its perceived value as a leading indicator has made understanding housing cycles important to economic forecasters.

Second, understanding investment behavior in the economy is important for understanding aggregate economic volatility. Investment spending is the most volatile component of aggregate demand, and spending on new home and apartment construction (residential fixed investment) is the most volatile component of total investment spending.[3] Thus, although residential investment is a smaller proportion of GNP than business fixed investment, its volatility has a particularly important influence on the volatility in national income over time.

Finally, because housing activity apparently has been sensitive to changes in interest rates, the housing sector historically has been an important channel through which monetary policy has influenced economic activity. In fact, there is some evidence that the economy may react more quickly to shocks that make their way through the housing component of aggregate demand than through the business fixed investment component.[4] Thus, when monetary authorities have decided to slow the national economy to bring inflation down, they typically have used monetary restraint to raise interest rates, which tended to contract the housing sector, and in this way, they were able to effect the desired cooling of the overall economy.[5]

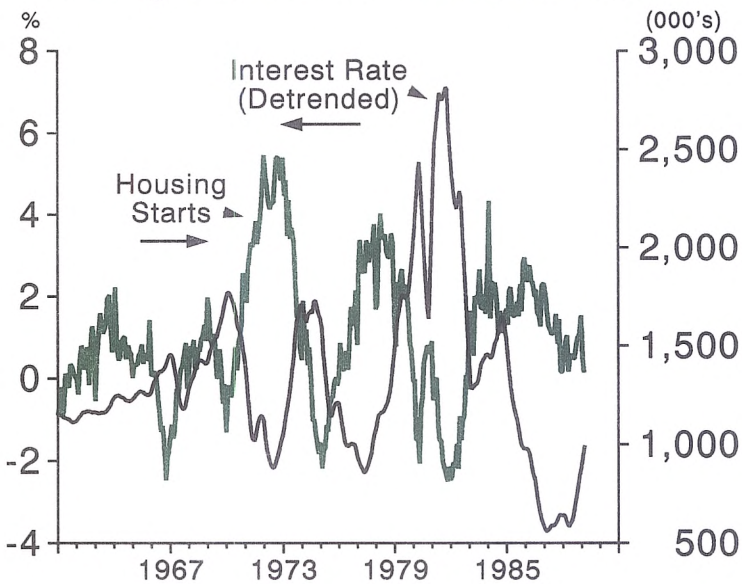
A key linkage in this process, of course, has been the relationship between housing activity and the level of interest rates. Historically there does appear, indeed, to have been a strong, inverse relationship between disturbances to interest rates and changes in housing starts. In recent years, however, the relationship between housing and disturbances in interest rates appears to have become less regular. (See Chart 1.)

This article explores the changes in the relationship between interest rates and housing, and tries to examine the potential causes of this changed relationship. The extensive financial innovation and reform that occurred in banking and mortgage markets in the early 1980s changed the way in which housing lenders raised their funds, and changed the types of mortgage instruments that were made available to homeowners. The empirical evidence presented in this paper suggests that the relationship between housing starts and interest rates changed significantly in the period after deregulation.

The remainder of this paper is structured as follows. In Section I, the various theoretical links between interest rates and housing are presented, and in Section II, the influence of financial reform on these links is hypothesized. As we shall see, there are numerous potential channels by which housing may have been affected by financial innovation and regulatory reform. In Section III, the relationships between interest rates and housing and between funds flows and interest rates are examined empirically. The data reveal that, in fact, a significant change in these relationships has occurred.

The article concludes (in Section IV) with a discussion of the policy implications of these findings. These implications relate both to the effects of the changes on monetary policy transmission and the likely effects on housing investment.

Chart 1
Housing Starts and Interest Rates



I. Interest Rates and Housing: Channels of Influence

Because of the intimacy of the relationship between housing and national income, it is important to understand the behavior of housing cycles, particularly as they relate to movements in interest rates, since the level of interest rates generally has been the control variable for monetary policy. In this section, I discuss the various linkages between housing investment demand and the interest rate that likely operated during the period prior to the recent financial reforms. To do so, I posit a simple model of housing investment.

A User Cost Model of Housing Investment

In this model, housing investment is a function of housing demand and a (lagged) supply process. Housing demand, or the desired stock of housing (H^*), is determined by the periodic costs of owning a unit of housing, that is, the user cost of housing, and demographic factors:

$$H^* = H^*(U, W, D) \qquad (1)$$

where

U = user cost
W = household wealth
D = various demographic factors, such as population, household formation behavior, etc.

The user cost of housing capital (or, sometimes more ambiguously, the rental price) represents all current, out-of-pocket costs and net foregone income that are associated with owning a home. A major component of the user cost is the interest cost associated with an investment in housing. In addition, costs are affected by (economic) depreciation of the structure, maintenance expenditures, and expected changes in the market price of the housing unit. This latter component, of course, can reduce or increase user costs depending on whether inflation or deflation in the market price of the unit is expected. Tax policy also affects user costs through the treatment of both interest expenses and capital gains in the tax code.

The relationship between the user cost, U, of a unit of housing, the real interest rate, r, and the other components of the user cost can be stated more precisely as:

$$U = P[(1-t)(i) - (1-c)(h) + d] \qquad (2)$$

where

P = market price of the housing unit
t = marginal tax rate on normal income
c = marginal tax rate on capital gains
d = rate of economic depreciation of the unit per period
i = the nominal interest rate = $(r + e)$
h = expected house price inflation
r = real interest rate
e = rate of inflation expectations

or, equivalently, if expected house price inflation, h, equals the expected rate of general price inflation, e,

$$\begin{aligned} U &= P[(1-t)(r+e) - (1-c)(e) + d] \\ &= P[(1-t)r - (t-c)e + d]. \end{aligned} \qquad (3)$$

From Equation (3), it is clear that user costs increase with the after-tax real interest rate, $(1-t)r$, and the rate of depreciation, d. User costs decline with increases in inflation expectations, e, as long as the tax treatment of capital gains is favored (that is, as long as $c < t$).

If the desired stock of housing equals the actual existing stock at the prevailing user cost (that is, if $H^* = \underline{H}$), then no investment in housing will occur. More likely, however, the desired stock is less than or greater than the actual, and investment (I) will occur as:

$$I = s(H^* - \underline{H}), \qquad (4)$$

where

s = an adjustment weight or function

and the actual housing stock will adjust toward the desired stock at a speed that depends upon the nature of the supply adjustment process, s. In markets in which the supply process is elastic, the adjustment in the housing stock will take place primarily through investment or disinvestment; less elastic supply conditions will result primarily in price adjustments to bring the desired and actual stocks into balance.

The Effects of Interest Rate Changes

We can now discuss the ways that interest rates affect housing investment. We begin first by discussing the channels of influence revealed by the simple user cost model described above. As we will see, however, in the period prior to recent financial market reforms, institutional factors created additional "channels" of influence.

Channel 1: Simple User Cost Effects

In the simple user cost model presented above, the interest rate is the key component of user costs, revealing a direct channel of influence of interest rates on housing. The precise effects of a change in interest rates depends, however, on whether the changes in interest changes occur because of changes in the underlying real rate or changes in inflation expectations.

Specifically, the effect on housing is unambiguous when a rise in the real interest rate occurs. As the rate, r, in Equation (3) rises, the user cost rises and depresses stock demand. This, in turn, results in depressed net new investment (manifested in reduced housing starts, for example), and possibly depressed real housing prices, depending on the elasticity of the housing supply response.

On the other hand, the effect of an increase in interest rates associated solely with an increase in inflation expectations is less clear. Equation (3) implies that if the tax rates on normal income and capital gains are the same, an equal increase in general and housing price inflation expectations would have no effect on housing demand. If, however, the capital gains tax rate is effectively lower than the normal tax rate (the case in the U.S.[6]), then an increase in inflation expectations can actually *lower* the user cost of housing capital because the higher interest costs are more than offset by the expected after-tax gain in the value of the housing assets. If the user cost is lower, then the demand for the housing stock increases, and with it, the price of housing and/or investment.

In summary, the demand for housing is a function of interest rates through their effect on the user cost of housing. The link between housing and interest rates is a negative one when the real interest rate changes and a positive one when inflation expectations change (assuming that capital gains receive preferential tax treatment, as is the case currently).

Channel 2: Credit Scoring and Affordability

A second channel by which interest rates potentially have influenced housing is the so-called affordability constraint, which arises out of the loan qualification process. This channel was a particular source of concern in the mid- to late-1970s. Because state laws limit lenders' ability to secure mortgage loans via the non-housing net worth and/or future income of households, the current home value and current income of the household play an important role in credit scoring or loan qualification standards. The standards usually are stated as limits on the ratio of mortgage payment size to household income, among other variables. To the extent that these standards do not change (because of constraints on mortgage design, or sluggish adjustment of the standard due to regulation or convention) they can become an additional source of influence for interest rate shocks. If interest rates rise abruptly, for example, the effective supply of mortgage credit to borrowers—under a fixed set of standards—can change abruptly as well.[7]

The affordability effect can be viewed in the context of the user cost model presented above as an additional factor that implicitly augments the interest rate component of user costs. To the extent that affordability constraints are binding, therefore, affordability constraints have the effect of depressing housing demand and lowering either housing investment activity or prices, or both.

To summarize, affordability considerations reinforce the depressing effect of higher real interest rates on housing investment and prices. Affordability considerations also will offset, at least partially, any stimulative effects of higher inflation expectations on housing investment. The extent to which affordability considerations depress housing investment and prices will depend upon the availability of ameliorative mortgage designs.

Channel 3: Interest Rates and Disintermediation

A third channel through which interest rates may have influenced housing investment is the phenomenon known as "disintermediation." Disintermediation refers to the tendency of funds to flow away from conventional housing lenders (such as thrifts and banks) when interest rates rise suddenly. Disintermediation occurred historically because the conventional mortgage intermediaries faced restrictions on their ability to pay deposit rates that were competitive with open-market investment opportunities. Faced with higher-yield opportunities elsewhere, consumers are said to have moved their deposits out of financial intermediaries, thereby reducing the funding available to conventional housing lenders. To the extent that other sources of mortgage funding were imperfect substitutes for financial intermediaries, the result was an increase in the cost of mortgage funds.

It is unlikely that the effects of disintermediation could have persisted for long periods of time, however. After the initial effects of an increase in interest rates, investment in mortgage debt by other lenders should have provided an offsetting supply of mortgage credit. For example, lending by non-depository intermediaries (such as insurance companies), issuance and sale of mortgage-backed securities by banks and thrifts, and financing by home sellers most likely increased to offset the decline in traditional mortgage intermediation.

In addition, even if banks and thrifts have been restricted in their ability to compete via higher deposit rates, in the long run, they were able to attract funds by increasing the services offered their depositors. This non-pecuniary form of competition eventually would have drawn some funds back to the affected banks and thrifts.

This argument suggests that disintermediation, to the extent it was influential, had primarily transient effects related to sudden *changes* in the level of interest rates and not to the level itself. The extent to which disintermediation affected the supply of mortgage funds depends on a number of factors, including the breadth and sophistication of the mortgage-backed securities market and the speed with which the deposit rate regulations could be circumvented by banks and thrifts. Disintermediation is cited as a significant contributor to short-term housing cycles in the 1960s and 1970s not only because deposit regulations were binding at times during this period, but also because the mortgage-backed securities market was not yet highly developed.

From a user-cost perspective, the disintermediation phenomenon would be manifested in price- or non-price rationing of mortgage funds that effectively would raise the interest cost component of housing user costs. The expected effect of this channel of influence, therefore, would be for higher interest rates (transiently, at least) to depress housing demand, housing investment, and housing prices.

II. The Effects of Financial Change

In the early 1980s, major changes in the financial system occurred that may have affected the functioning of these various linkages. Legislative reforms and financial innovation affected mortgage instrumentation, mortgage intermediaries, and household mortgage demand.

A Changing Marketplace

There were both legislative and market changes in the early 1980s that may have influenced the interest rate/housing link. Two key pieces of legislation affecting mortgage markets that were enacted in the early 1980s were particularly important. The first, the Depository Institutions Deregulation and Monetary Control Act (DIDMCA), was passed in 1980. Title II of this Act provided for interest rate ceilings on time and saving deposits at banks and thrifts to be phased out over a six year period. Title III provided nationwide authorization of interest-bearing transactions accounts as of January 1, 1981. These accounts were negotiated order of withdrawal, or NOW, accounts with a regulated maximum rate.

Because depository institutions were believed to still be at a disadvantage vis à vis the continued, intense competition from (non depository) money market mutual funds, a second piece of legislation was passed in 1982. The Garn-St. Germain Depository Institutions Act of 1982 authorized (in its Title III) the money market deposit account (MMDA). The MMDA required a minimum balance and had restricted transactions capability, but offered an unregulated deposit rate. It was widely available by the end of 1982. On January 5, 1983, so-called Super Now accounts, with unregulated deposit rates, were permitted. Thus, effective deregulation of retail transactions deposit rates occurred sometime between 1980 and 1983, although full removal of rate ceilings and account minimums on all types of retail accounts did not occur until 1986.

These Acts had features that also affected the mortgage markets. Title V of DIDMCA, for example, authorized an override of state usury provisions on loans secured by liens on eligible residential real estate and made after March 31, 1980. Title II of the Garn-St. Germain Act, also preempted state-imposed restrictions on the execution of the due-on-sale clause in mortgage contracts.[8]

In addition to legislative changes, several important regulatory changes broadened the types of mortgage instruments that could be offered by banks and thrifts. First, in 1981 the Federal Home Loan Bank Board permitted thrifts to offer adjustable rate mortgages on a widespread basis.[9] Since regulation and the conventions of the secondary markets have strongly influenced the types of mortgages that lenders have issued ever since the 1930s, this change in regulation means that prior to 1981 it is likely that well over 95 percent of all residential mortgages issued were conventional, fixed-rate, self-amortizing instruments, although accurate statistics to support this observation unfortunately are not available.

In addition, the technology of the mortgage marketplace was changing during the period of the early 1980s. As a result of the continued development of the secondary mortgage market, in particular, newly-originated mortgages no longer needed to be funded within the bank or thrift portfolio. Instead, mortgages could be used to create mortgage-backed securities which could then be sold to a variety of institutional and private investors. This process, known as securitization, was facilitated by government-backed mortgage agencies which provided credit enhancement in the form of principal and interest guarantees to investors in the securities. Development of the secondary mortgage market was particularly rapid in the early 1980s. The volume of contracted mortgage commitments of the Federal Home Loan Mortgage Corporation (FHLMC), for example, grew from about $7 billion in 1981 to almost $33 billion in 1983.

Effects on the Housing Market

These legislative reforms and market developments had the effect of facilitating better matching of the needs of demanders and suppliers of mortgage credit over the interest rate cycle, and as a result, affected most of the interest rate/housing channels discussed earlier. First, the deregulation of deposit rates removed the primary cause of financial disintermediation. Deposit rate flexibility enabled banks and thrifts to price their deposits more competitively with non-deposit investments. Thus, when general interest rates rise in the current regulatory environment, there need be no tendency for depositor funds to flow out of financial intermediaries into investments in the primary securities markets.[10] This presumably has had the effect of making mortgage supply less cyclically-sensitive.

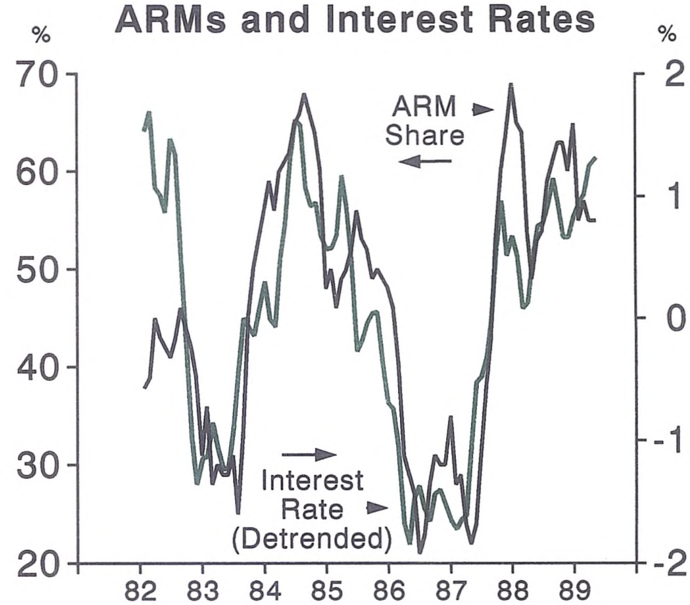
Second, the continued development of secondary mortgage markets also helped to make mortgage supply less cyclical. A mortgage lender having difficulty attracting funds can now originate a qualified mortgage, and sell it into a very liquid secondary market.

Third, the availability of the adjustable rate mortgage after April 1981 likely affected the channel relating to the affordability constraint. The ARM generally has initial payments that are lower than those of fixed rate mortgages, making any given payment-to-income test less binding. In addition, the flexibility of the instrument's payment structure (allowing such things as low buy-in or "teaser" rates, negative amortization, and so on) can be used to tailor the instrument to borrower needs over the business cycle. This, too, should have the effect of better insulating mortgage supply conditions from interest rate cycles.

The adjustable rate mortgage may be influential in another way, as well. It can be shown that borrowers may prefer ARMs over fixed rate instruments when they believe that their own income is likely to fluctuate with future interest rate movements. Elimination in 1981 of restrictions on ARMs therefore likely made the mortgage market more "complete"; that is, it may now be better able to efficiently match borrowers' and lenders' needs as they vary over the business cycle, reducing the cyclical linkage among interest rates, mortgage credit, and housing investment.[11]

As Chart 2 illustrates, the share of ARMs in new mortgages varies with the interest rate cycle. The chart plots the deviations in interest rates from a simple linear time trend, and the share of new mortgages issued as adjustable rate instruments. When interest rates rise above their trend, the ARM share also rises, consistent with the notion that the ARM instrument does, indeed, help to "buffer" somewhat the effects of interest rate spikes.

Chart 2
ARMs and Interest Rates



In summary, there are a number of reasons to expect a weakening of the linkage between interest rates and housing investment activity at some point in the early 1980s. The linkage likely has weakened along with the weakening of the secondary channels of influence—affordability and credit scoring constraints, financial disintermediation effects, and other mortgage instrumentation constraints.

III. The Interest Rate Link: The Empirical Record

In this section, the available data are examined to determine whether the changes in the early 1980s actually diminished the strength of the relationship between interest rates and housing. The empirical approach employs simple, time-series models estimated using data from the period from 1960 to 1989. In this section, I test for changes in the effect of interest rates on housing starts and for changes in the effect of interest rates on fund flows to institutional housing lenders (thrifts and banks). The model presented above permits analysis of the effect on housing of one-time disturbances or "shocks" to interest rates and comparison of the size of this effect in the pre- and post-deregulation periods. My interest in testing the effects on housing starts follows directly from the hypothesis that positive disturbances to interest rates typically have depressed housing demand and, thereby, housing investment. My interest in studying the effect of interest rates on fund flows is to determine whether mortgage-supply phenomena were, in fact, a channel of influence.

The Basic Model

The models estimated in this section use simple, vector autoregression (VAR) systems. These VAR models employ generalized time-series equations to identify dynamic relationships among the variables of interest. Such models are particularly appropriate in this application because they permit exploration of shocks and other dynamic interactions among the variables. In addition, their atheoretical structure is desirable in this context because we are interested in changes in linkages across periods, rather than testing a particular model specification.

The estimated VAR equation systems involve the current value of each variable of interest regressed on lagged values of itself and every other variable in the system. In a two variable (X and Y) system, for example, the estimated equations would be:

$$X_t = c_1 + \sum_{i=1}^{n} a_{1i} X_{t-1} + \sum_{i=1}^{n} b_{1i} Y_{t-1} + e_{1t} \qquad (5)$$

$$Y_t = c_2 + \sum_{i=1}^{n} a_{2i} X_{t-1} + \sum_{i=1}^{n} b_{2i} Y_{t-1} + e_{2t} \qquad (6)$$

where a, b, and c are estimated coefficients and n is the lag length employed.

In the empirical work below, equations similar to those in (5) and (6) are estimated using monthly data on interest rates and various housing-related variables. A comparison of the relationships estimated for the period prior to deregulation with those for the period afterward can detect changes in the interest rate/housing linkage. For the purposes of the analysis below, the pre-deregulation period is assumed to span from 1960 to 1982, and the post-deregulation period extends from 1983 to 1989.[12]

The linkages between the interest rate variables on the one hand, and housing starts or funds flows, on the other, are explored in a series of simple VARs, rather than in one large VAR system, which would take into account all the interrelationships among these variables. As a practical matter, the paucity of data in the post-deregulation period constrains the size of the VARs that may be employed. Thus, housing starts and funds flows are studied one at a time, paired with the interest rate variable(s).[13]

Several statistical tests are presented to demonstrate the changes in the interest rate/housing linkage. First, a Chow test is used to compare the VAR systems for the pre-deregulation period with those estimated over the subsequent period to determine whether the estimated equations differ significantly between the two estimation periods.[14]

Second, the share of the total observed variation in housing starts "explained" by interest rate variation is examined for the two periods. While such "variance decomposition" exercises cannot reveal changes in statistical "causality" with great precision, changes in the contribution of interest rates to the variance in starts between the two periods is suggestive of a change in the underlying structure of the housing market.[15]

Finally, impulse response functions are estimated and presented for both the pre- and post-deregulation periods. These relationships project the effects into future periods of an hypothetical, one standard deviation shock in interest rates. Unlike a simple comparison of coefficients, the impulse response functions incorporate all direct and feedback effects. They provide a graphical summary of interest rate effects before and after deregulation.

The Data

The models reported in this paper are all two- or three-variable VARs that use monthly data and a 12-period lag structure on all variables. The interest rate (TBILLS) variable employed is the short-term interest rate, measured by the 90-day Treasury bill yield. In addition, in some of the VARs reported below, the difference between the short- and long-term interest rate (LNGMSHRT) is also included as a proxy for the effects on the yield curve that would accompany a change in long-term inflation expectations. As discussed in Section II, whether an increase in interest rates is due to the real rate or inflation expectations may have different effects. The long rate used is the AAA corporate bond yield.[16]

Housing starts data (rather than data on housing investment) are employed in this study to permit a monthly time frame for analysis and to avoid the arbitrary valuation assumptions that must be made to calculate housing investment flows. The housing starts variable (STARTS) is seasonally adjusted. The aggregate of funds flows to thrifts and commercial banks (FUNDTOT) is measured as net changes in total share balances and total deposits.

Empirical Results

The results presented first shed light on the hypothesis that changes in mortgage instrumentation and other factors have relaxed the "affordability" constraint, thereby reducing the "direct" effect of interest rate cycles on housing starts.

Housing Starts and Interest Rates

Housing start relationships are studied using two simple VAR systems, compared over two periods. In the first VAR, the only interest rate variable is the Treasury bill rate. In the second, both the Treasury bill rate and the difference between the long rate and the short rate are used.

Chow tests suggest that in both VARs, the results from the two periods are significantly different at a confidence level of over 90 percent. In addition, the pattern of coefficients (for brevity, not shown here) in both models suggests that the relationship between interest rates and housing starts has changed significantly since 1983.

Table 1 presents the variance decomposition obtained from the two models for both time periods. As would be expected if the interest rate channel had weakened in the later period, the contribution of interest rates to the total variation in housing starts has declined.

Table 1

Interest Rate Variation and Housing Starts Variation*

Percent of STARTS variance explained by:	VAR 1		VAR 2	
	Pre-83	Post-82	Pre-83	Post-82
STARTS	39.6	78.5	32.9	43.9
TBILLS	60.4	21.5	56.9	28.2
LNGMSHRT	—	—	10.2	27.8

*Variance decomposition of STARTS, from 12-lag VARs; in percent of variation explained, measured at 24 months

The impulse response functions in the three panels of Chart 3 reveal the extent of the changes in the relationship more graphically by tracing out the effects on housing starts of a positive, one standard deviation shock to the interest rate variables. Chart 3A presents the impulse response of housing starts to a one-standard deviation shock in the Treasury bill rate, as estimated from the first model. Panels B and C of Chart 3 display the same thing for the two variables of the second model. These graphs illustrate that whereas prior to deregulation, a one

Chart 3B
Impulse Response of Housing Starts to Shock in Treasury Bills

(VAR Includes STARTS, TBILLS, LNGMSHRT)

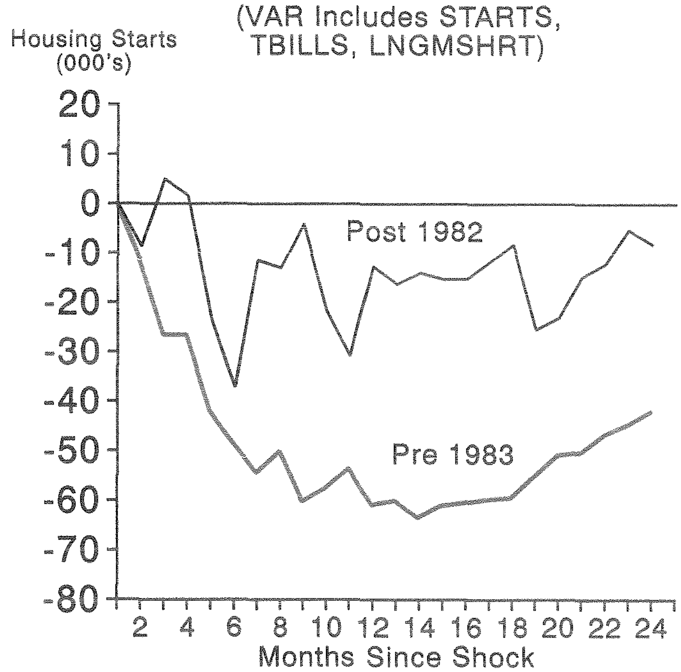


Chart 3A
Impulse Response of Housing Starts to Shock in Treasury Bills

(VAR Includes STARTS, TBILLS)

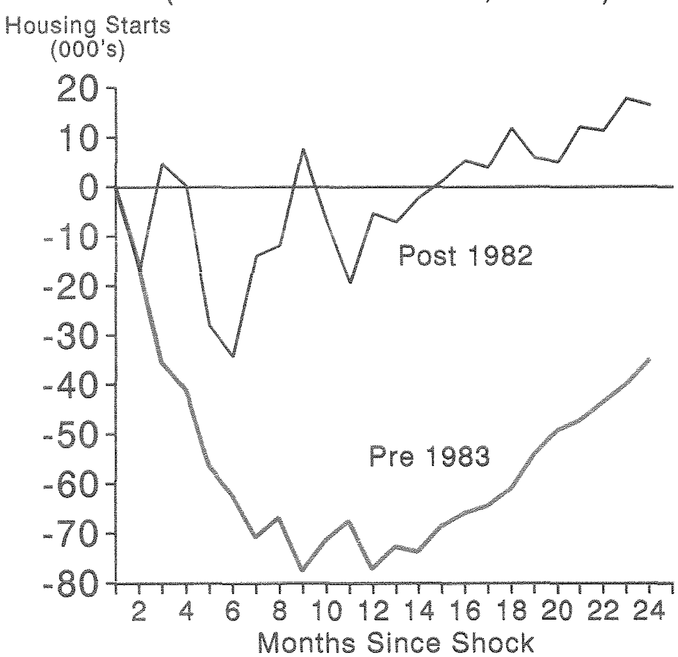
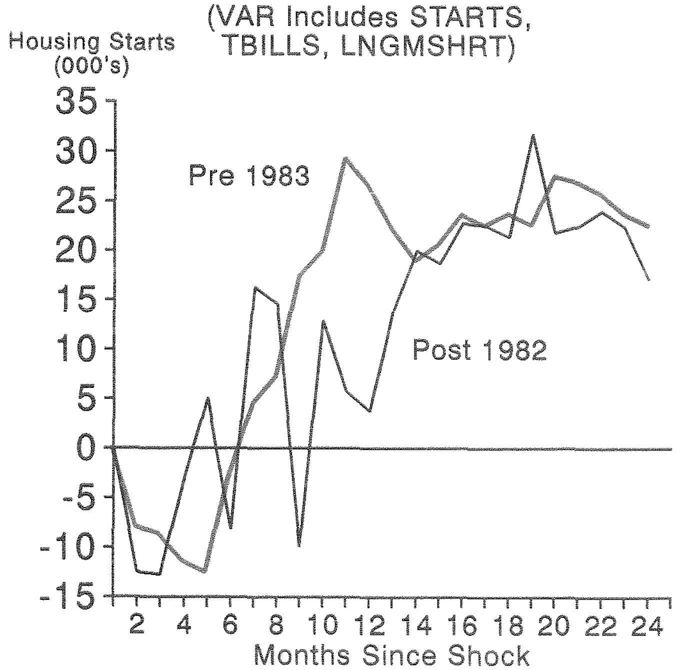


Chart 3C
Impulse Response of Housing Starts to Shock in Yield Curve Tilt

(VAR Includes STARTS, TBILLS, LNGMSHRT)



standard-deviation shock to interest rates would have depressed housing starts by up to 80 thousand units, such an effect is largely absent in the period after 1982.

The second formulation of the VAR can be used to assess whether there has been a change in the effect on starts of an upward tilt in the yield curve. Such an effect, which might be associated with increased long-term inflation expectations, is represented by LNGMSHRT.

As is apparent from the variance decomposition results in Table 1, in both periods, LNGMSHRT explains a sizable proportion of the variation in starts. From the graphs of the impulse response functions, it is apparent that in both periods, the effect of an upward tilt in the yield curve is to stimulate housing starts. This means that the capital gains effects of higher long-term inflation expectations apparently ultimately offset any affordability effects of higher long-term interest rates. Prior to 1983, however, the effect reached a peak only after 18 months; in the period afterward, the positive effects occur approximately six months earlier, and are larger on average. The contribution of LNGMSHRT to explaining the variance of starts also is larger in the later period. This is consistent with the notion that improved mortgage instrumentation in the later period resolves the affordability problems associated with higher rates more quickly and easily than was the case in the earlier period.

These results confirm that the relationship between housing starts and interest rates is different in the post-1982 period from the pre-deregulation period.

Funds Flows and Interest Rates

The fact that housing starts appear less sensitive to interest rates after 1982 than before does not establish which of the various channels of influence has changed to produce this effect. Thus, the second hypothesis I will test is that deregulation of deposit rates and continued development of alternative mortgage finance mechanisms have made housing less sensitive to the effects of "disintermediation." This hypothesis has two parts: (1) that funds flows into banks and thrifts are now less sensitive to interest rates, and (2) that the financing of housing starts has become generally less affected by patterns of funds flows to traditional lenders.

If deposit rate deregulation, and the subsequent decline of the disintermediation phenomenon are important, the funds flow variable should be sensitive to interest rates prior to 1983, but less so afterward. Additionally, if deregulation and growth of the secondary market has made housing investment less tied to funding from traditional intermediaries, any effect that fund flows have on housing

Table 2
Interest Rate Variation, Funds Flows and Housing*

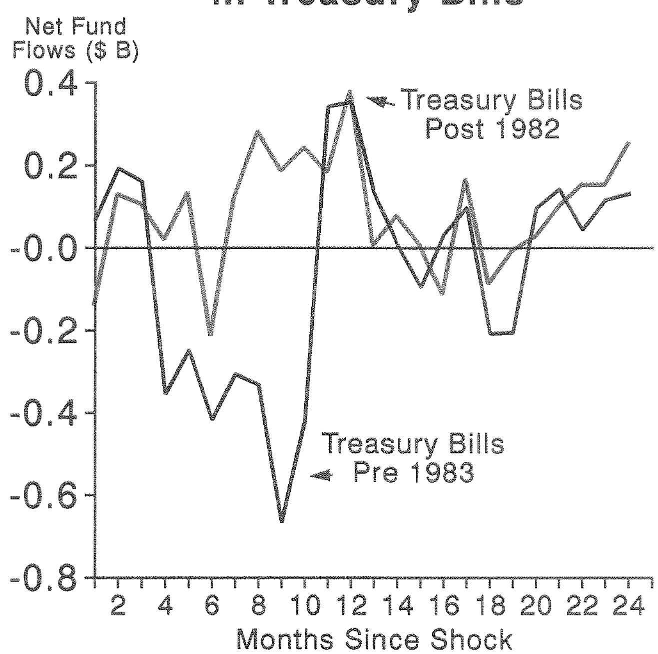
Percent of variance explained by variables below:	VAR 1 FUNDTOT Pre-83	VAR 1 FUNDTOT Post-82	VAR 2 STARTS Pre-83	VAR 2 STARTS Post-82
STARTS	—	—	34.1	65.8
TBILLS	88.8	69.1	64.6	25.8
FUNDTOT	16.2	30.9	1.3	8.4

*Variance decomposition from 12-lag VARs; in percent of variance explained, measured at 24 months

starts in the pre-1983 period should weaken in the period afterward. This is tested with VARs that include the total of funds flows to banks and thrifts (FUNDTOT). To test whether interest rate channels other than deregulation play a role, the interest rate variable is included in these VARs as well.

In Table 2, and Charts 4 and 5, the results are presented of VARs involving housing starts, fund flows, and the interest rate for the two sample periods. The variance decomposition results in the first two columns of Table 2

Chart 4
Impulse Response of Net Fund Flows to Shock in Treasury Bills



Note: Post 1983 normalized to pre 1983 levels for comparison purposes.

are consistent with the hypothesis that disintermediation has declined in the sense that fund flows are less sensitive to interest rate shocks in the post-82 period. This is confirmed more formally by a Chow test which finds the estimated relationship for the two periods to be significantly different at better than the 90 percent level.

The impulse response functions graphed in Chart 4 depict the effects of these changes. Prior to 1983, a one standard deviation increase in the Treasury bill rate resulted in a decline in net fund flows to banks and thrifts that began about three months after the shock, and extended for seven or eight months. In the period after 1982, an effect of this scale appears to be absent.

In the third and fourth columns of Table 2, the variance decomposition results from the VARs linking housing starts to fund flows and interest rates are reported. In this case, the impact on housing starts of a shock to interest rates declines as expected. The impact of fund flows increases, suggesting that traditional intermediaries, if anything, play a more important role in the post-82 period. However, this may simply be a result of higher variance in fund flows after 1982. For this reason, it is important to inspect the impulse response function. The change in the coefficients estimated in the two periods is significant here as well, as confirmed by a Chow test.

In panels A and B of Chart 5, the effects of positive shocks to fund flows and Treasury bills are depicted for the pre-1983 and post-1982 periods, respectively. As in the simpler VARs discussed above, a depressing effect on housing starts of shocks to the Treasury bill rate is observed in the pre-1983 period. A one-standard deviation shock results in a 70 thousand unit decline in housing starts. In this model, the additional funds flow variable also has an effect on starts; a shock to fund flows does appear to stimulate starts, suggesting that during this period housing was linked to the funding capability of the traditional housing lenders.

Panel B of Chart 5 suggests that both effects are much less pronounced in the post-1983 period. This finding is consistent with the notion that not only is housing less sensitive to interest rates directly, but also that the supply of funds to housing from other sources (via the mortgage securities market, for example) has increased.

Chart 5A
Impulse Response of Housing Starts to Shocks in Treasury Bills & Net Fund Flows

(Pre 1983)

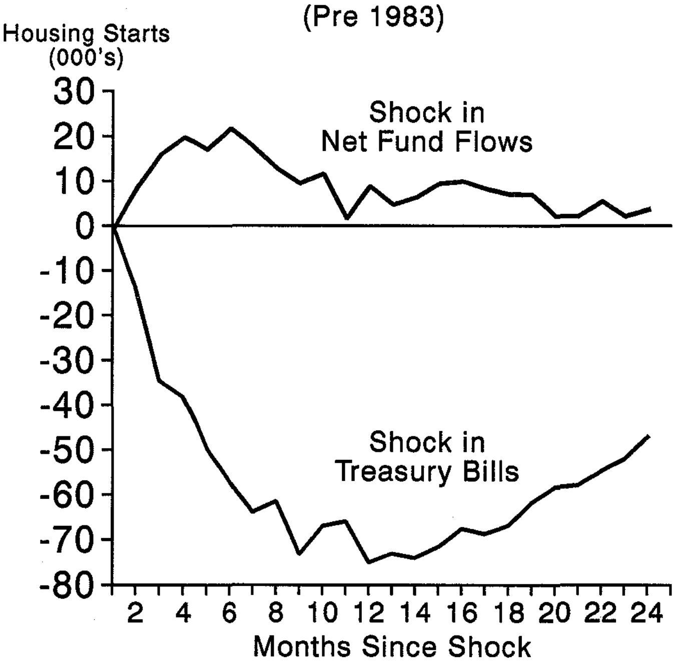
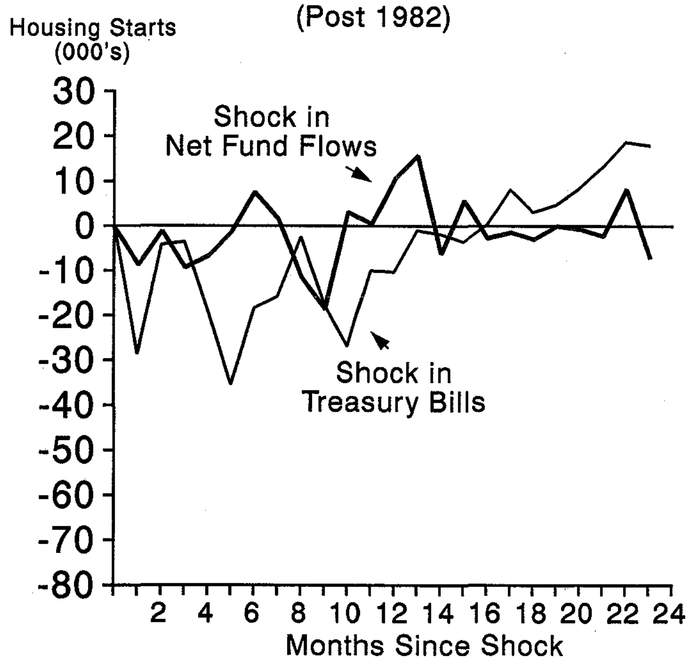


Chart 5B
Impulse Response of Housing Starts to Shocks in Treasury Bills & Net Fund Flows

(Post 1982)



IV. Conclusions and Policy Implications

The potency of the interest rate/housing linkage appears to have changed significantly in the period following extensive financial deregulation in the early 1980s. In this paper, simple time series statistical models were used to measure the changes in the strength of this linkage, and to explore the possible causes of the changes.

The available data allow demonstration of a strong association in time between the changes in this linkage and changes in the regulation of mortgage lending institutions. The data also allow testing of the independent causal linkage between housing starts and fund flows into traditional mortgage lenders. The results, therefore, are consistent with linkages associated with disintermediation processes, affordability constraints, mortgage instrumentation restrictions, and growth of secondary markets.

Less important than the precise linkage, however, is the fact that the strength of the link appears to have weakened considerably in the post-82 period. This change has the greatest import for investors, builders and owners of housing in the United States, since it means that the housing sector is less likely to be buffeted severely during periods of economic policy manipulation.

The housing market appears to be functioning at a level of housing production of about 1.6 million units per year; this puts the market in a steadier environment that is below the peaks of earlier cycles, but is also well above the troughs. In such an environment, participants in the housing market can direct their resources to responding to other planning parameters, such as local economic and demographic conditions.

The weakening of the interest rate link to the housing cycle also may be important from the standpoint of the conduct of national economic policy. With a weaker link between interest rates and housing, national output levels become less sensitive to interest rate disturbances, at least on a cyclical basis. This is a desirable prospect, of course, to those economists who would prefer to see the economy buffered against most macroeconomic disturbances, which they view as originating from the mismanagement of monetary aggregates and interest rates. If, on the other hand, manipulation of aggregate real output levels is a key element of effective national economic policy, the increased insulation of the housing sector may make such management more difficult.

NOTES

1. For example, the standard deviation of housing starts from its trend from 1962 to the present has been one-fifth of its mean level, versus only one-twentieth for real GNP.

2. R. E. Hall and J. B. Taylor, *Macroeconomics* (New York: W. W. Norton & Company), 1986, p. 203.

3. The other two components are nonresidential fixed investment (purchases of new plant and equipment by businesses) and inventory investment (changes in stocks of goods produced but not yet purchased).

4. See P. K. Clark, "Investment in the 1970s: Theory, Performance and Prediction," *Brookings Papers on Economic Activity*, Washington, D.C., 1979, pp. 73-113.

5. The role of residential investment in business cycles is discussed along these lines in Dornbusch and Fischer, *Macroeconomics*, New York: McGraw-Hill, 1987, pp. 317-326.

6. There are several aspects of U.S. tax policy that make the capital gains tax rate less than the tax rate on "ordinary" income. First, the statutory rate on realized, long-term capital gains historically has been lower than the rate on ordinary income. Thus, although the tax reforms of 1986 made these rates the same, over most of the period covered in this paper, there was a significantly lower long-term capital gains rate. Second, capital gains generally have been taxed in the U.S. only as they are realized (rather than on an "accrual" basis). This ability to time and delay capital gains tax burdens (but not ordinary income tax burdens) is an additional source of preferential treatment of capital gains. Finally, housing capital gains have enjoyed an additional advantage in that the tax burden can be sheltered beyond the date of realization if the proceeds from the sale of one primary residence are rolled over into another home within a specified time (presently, 2 years).

7. The size and significance of the disintermediation phenomenon was the subject of considerable discussion. See, for example, F. Arcelus and A. Meltzer, "The Markets for Housing and Housing Services," *Journal of Money, Credit and Banking* (1973), pp. 78-99, and D. Jaffee and K. Rosen, "Estimates of the Effectiveness of Stabilization Policies for the Mortgage and Housing Markets," *Journal of Finance* (1978), pp. 933-46.

8. The "due-on-sale" clause gives the lender the option to terminate the loan secured by a home when the home is sold.

9. The Federal Home Loan Bank Board in April 1981 allowed thrifts under its supervision to offer ARMs. This power was extended to other institutions as the result of a provision of the Garn-St. Germain Depository Institutions Act, passed in 1982.

10. Whether the relationship between interest rates and deposit flows disappears altogether, however, is less certain. Depending upon the deposit pricing strategy of banks and thrifts, and how rapidly and completely open-market rates are matched, there still may be some reaction of deposit flows to interest rate changes. See M. J. Flannery, "Retail Bank Deposits as Quasi-Fixed Factors of Production," *American Economic Review*, June 1982, pp. 527-536, for a discussion of the process that might cause banks to make less-than-complete adjustments to open-market rates.

11. See, for example, H. R. Varian, "Divergence of Opinion in Complete Markets: A Note," *Journal of Finance* (1985), pp. 309-317.

12. The selection of the date that constitutes the break between the pre- and post-deregulation periods necessarily is somewhat arbitrary. Conceptually, it should be possible to find a breakpoint that maximizes the differences between the pre- and post-period VAR estimates. As a practical matter, the results are relatively insensitive to a range of breakpoints a year or so on either side of the chosen date.

13. The alternative of estimating the VARs with exogenous period dummy variables, and interactions of those dummies with the interest rate variable(s), also was explored. This permits a much larger model, but that structure complicates testing of the significance of changes in the model's coefficients. In addition, sample size limitations do not permit incorporating the full number of interaction terms in the larger model. Qualitatively, however, the findings are the same in either modelling context. The simpler methodology facilitates presentation of the findings.

14. The Chow test employs residual sum of squares (*RSS*) information from a regression spanning the entire data sample (*RSS*1), the first subperiod (*RSS*2), and the second subperiod (*RSS*3). An *F* test is then constructed as:

$$F = [(RSS1 - RSS2 - RSS3)/k] \,/\, [(RSS2 + RSS3)/(N1 + N2 - 2k)]$$

with degrees of freedom = $\{k, N1 + N2 - 2k\}$ where $N1$ is the sample size of the first subperiod, $N2$ is the sample size of the second subperiod, and k is the number of estimated parameters.

15. The use of the variance decomposition in this manner has two potential problems. First, a problem in interpretation of the variance decomposition can occur if the variance of interest rates changes significantly between the two periods. In such a case, the contribution of interest rates to the variance of housing starts may appear to have changed, but the measured effect is caused simply by the change in the variance of interest rates. Second, the "ordering" of the variables (which affects the precedence of shocks) can affect the results. In our case, however, we are interested only in comparisons across periods (not the levels of variance contributions per se). This is less affected by ordering considerations.

16. The mortgage rate is not used specifically because of the potential problems interpreting this series, as mortgage instrumentation and other features of the mortgage market change over the time frame of the analysis.

What Does Unemployment Tell Us About Future Inflation?

John P. Judd and Bharat Trehan

Vice President and Associate Director of Research, and Senior Economist, Federal Reserve Bank of San Francisco. The authors would like to thank Jack H. Beebe and the members of the editorial committee, Chan G. Huh, Brian Motley, and Carl Walsh, for helpful comments.

The unemployment rate commonly is used as an indicator of future inflation, with a low unemployment rate, for example, assumed to imply higher inflation. This negative correlation between current *unemployment and* future *inflation assumes that aggregate demand factors primarily are responsible for movements in these variables. However, as discussed in this paper, changes in aggregate supply conditions, such as technology and labor supply, also cause movements in unemployment and inflation. These factors lead to a* positive *relationship between future unemployment and current inflation. Consequently, a given rate of unemployment could be associated with almost any rate of inflation, depending on the source of the shock. In this paper, we attempt to disentangle demand and supply shocks, and analyze their influence on unemployment and inflation in the post-World War II U.S. economy.*

The level of the unemployment rate commonly is used as an indicator of future inflation. When unemployment is judged to be below (above) its long-run, or "natural" rate, inflation is projected to rise (fall) in the future. This negative correlation between unemployment and inflation is fundamental to the Keynesian, expectations-augmented Phillips curve, which expresses inflation as a function of the unemployment rate relative to its long-run level, expected inflation, and changes in certain relative prices such as those of oil and the dollar.

This interpretation of the relationship between unemployment and inflation focuses primarily on the effects of demand factors, such as monetary and fiscal policies. In recent years, however, macroeconomic research increasingly has incorporated such aggregate supply factors as changes in technology and the supply of labor in models of the behavior of the economy. "Real business cycle" models, in particular, attempt systematically to incorporate the effects of supply factors. As discussed below, this real business cycle approach[1] suggests a *positive* correlation between unemployment and inflation.

Conceptually, the correlations between inflation and unemployment implied by both the Phillips-curve and real business cycle models may coexist. The observed relationship in any given period thus depends on whether demand or supply factors were the more influential during that period. Accordingly, in this paper we estimate a model that treats unemployment and inflation as endogenous variables that respond to both aggregate demand and aggregate supply factors.

We find that both kinds of shocks are important in explaining movements in inflation and unemployment, and that both produce the well-known clockwise temporal loops observed when the actual inflation rate is plotted against the actual unemployment rate. Thus these loops, which commonly are presented in macroeconomic textbooks as arising from demand shocks in the context of the Phillips-curve relationship, also are consistent with supply shocks playing an important role. In addition, we find that while the effects of demand shocks on the unemployment rate reverse themselves in one to two years, the effects of supply shocks last much longer, and appear to have been responsible for large, persistent movements in the unemployment rate. Finally, the fact that *both* demand and supply shocks play significant roles diminishes the usefulness of the unemployment rate as an indicator of future inflation, since policy makers must be able to identify the source of a change in the unemployment rate before that variable can be used to make a forecast of future inflation.

The remainder of the paper is organized as follows. Section I spells out the kinds of correlations between inflation and unemployment that may be expected on *a priori* grounds in response to demand and supply shocks. Section II discusses the econometric method we use to estimate demand and supply shocks, and presents some evidence on the characteristics of these shocks. In Section III we present empirical estimates of how inflation and unemployment react to these shocks, and also the role played by the shocks in generating the observed correlation between inflation and unemployment over the past 25 years. Finally, Section IV discusses some policy implications.

I. Inflation-Unemployment Correlations in Theory

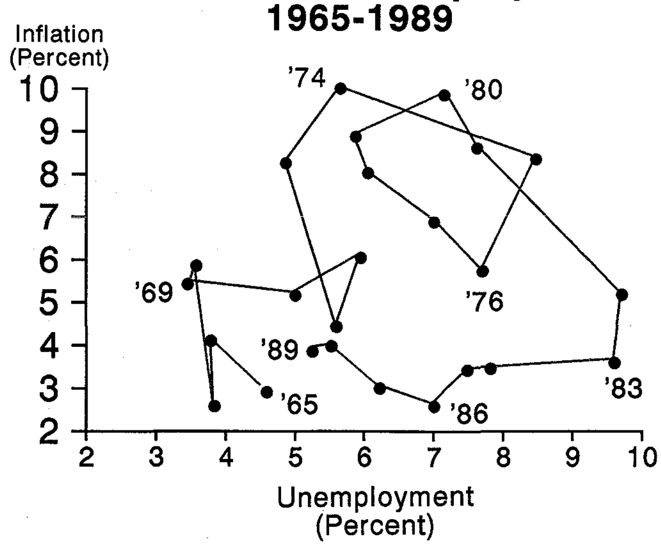
Keynesian theory stresses the role of aggregate demand factors in causing business cycles, and focuses on capacity bottlenecks caused by excess demand as the catalyst for inflation. In this view, prices rise to relieve shortages of labor and capital, and when monetary policy accommodates these price pressures, inflation results. The expectations-augmented Phillips curve embodies this hypothesis; in this case, for a given level of expected inflation, the difference between the prevailing rate of unemployment and the bench-mark rate (the so-called "natural" rate) provides a measure of aggregate-demand pressures on inflation.[2]

The inflation process can be illustrated with the following example. A positive demand shock induces firms to hire more workers. As the unemployment rate falls below the natural rate, labor markets become increasingly tight, and firms push wages up as they bid for labor. Faced with rising labor costs, firms raise product prices to maintain their mark-up over cost.[3] Thus, a decrease in the unemployment rate is followed by higher inflation. Ultimately, however, the unemployment rate returns to its original (long-run) level.[4] When the unemployment rate is graphed against the inflation rate over time, this sequence of events leads to clockwise loops similar to those shown in Chart 1.

Current research on the Phillips curve relationship also allows some kinds of relative prices to affect the inflation rate, such as changes in the relative price of oil and the real foreign-exchange value of the dollar. However, the Phillips curve captures only the direct price effects of a change in relative prices. By construction, it excludes possible effects on the unemployment rate of supply shocks associated with changes in relative prices. For example, a rise in the price of oil not only can be expected to raise the aggregate price index, but also may raise the unemployment rate.[5] Moreover, the Phillips curve omits by construction a broad range of other types of supply shocks, the most significant of which may be changes in technology and in the labor-leisure decisions of households. Attempts to incorporate aggregate supply factors into the Phillips curve model have been *ad hoc,* and do not represent a comprehensive and systematic treatment of this aspect of economic behavior.[6]

By contrast, real business cycle models attempt to explain business cycles entirely on the basis of real developments, such as shocks to labor supply and technology.[7] These models de-emphasize demand factors in much the same way that Keynesian models de-emphasize supply factors. The real business cycle approach received considerable impetus from the observation that the levels of many real variables, including real GNP, contain permanent, random-walk-like components.[8] Given that

Chart 1
Inflation and Unemployment
1965-1989



economic theory suggests that demand factors cannot permanently affect the levels of real variables, supply shocks, which can have permanent effects, must play a role in explaining fluctuations in real GNP.

The simple model of aggregate demand and supply commonly used in macroeconomic textbooks can be used to illustrate how a supply shock would affect the correlation between inflation and output. A positive technology shock, for instance, leads to an increase in aggregate supply (that is, a rightward shift in the aggregate supply curve), implying an increase in equilibrium output and a fall in the price level. In a dynamic context, a rise in aggregate supply would translate into lower inflation and higher real GNP growth.

Rigorously-derived real business cycle models that allow a role for money also predict such a negative correlation between inflation and output.[9] Moreover, a positive (negative) technology shock will have sustained negative (positive) effects on the unemployment rate in these models as long as searching for labor is costly.[10] For example, a higher marginal product of labor (positive technology shock) raises the marginal net benefit to searching for labor, and thus lowers the unemployment rate.

Putting these elements together implies a *positive* relationship between inflation and the unemployment rate. Whether the predicted co-movements are consistent with the clockwise temporal loops shown in Chart 1 depends upon the dynamic properties of the responses of inflation and unemployment to supply shocks. For instance, clockwise loops likely would result if the inflation rate responded first to supply shocks, and the unemployment rate responded afterwards. Theory does not predict the exact dynamic pattern that will occur, so the question must be resolved empirically.

Real business cycle and Phillips curve models both imply extreme views of the source of observed co-movements in the inflation and unemployment rate data. Theory does not rule out the possibility that *both* demand and supply factors operate simultaneously, and combine to produce the data we observe. The magnitude of each factor's independent influence, then, is an empirical issue. A balanced approach, in which neither demand nor supply factors are excluded, appears to be the most fruitful strategy for research. In the next section, we use an approach that is theoretically agnostic about the relative importance of demand and supply factors, and instead uses the data to estimate the magnitudes of each of those influences.

II. Estimates of Demand and Supply Shocks

In assessing the major forces determining the inflation-unemployment relationship, economic time series that directly measure aggregate demand and supply shocks would be most helpful. However, this is not possible in most cases.

Demand shocks can arise from a number of sources including changes in monetary policy, fiscal policy, inflation expectations, and consumer tastes, among others. Deregulation of the financial system has made the money supply (historically a good source of information on demand shocks emanating from Federal Reserve policy) a poor measure of these shocks.[11] Interest rates might provide an alternative measure since they are influenced by Federal Reserve actions. But because they also are influenced by other factors such as fiscal policy, inflation expectations, and aggregate supply, they are likely to be poor measures of monetary policy shocks as well.

With respect to variables representing fiscal policy, there are major problems in the national income accounts that make it difficult to obtain conceptually appropriate measures of government activity, including the inability to distinguish between capital and current expenditures and the exclusion of the "revenues" generated by the inflation "tax".[12] Other factors that can induce demand shocks—such as changes in the public's expectations of inflation or consumer confidence—also are difficult to measure directly.

Similar problems exist in attempting to measure supply shocks. These shocks originate from a variety of sources, including the development of new products (for example, computers), new ways to combine labor and capital more efficiently, changes in individuals' willingness to work, changes in tax laws, as well as sudden changes in the relative prices of important inputs to production such as oil. While certain taxes and relative prices can be measured directly, the other potential sources of supply shocks do not have direct empirical counterparts.[13]

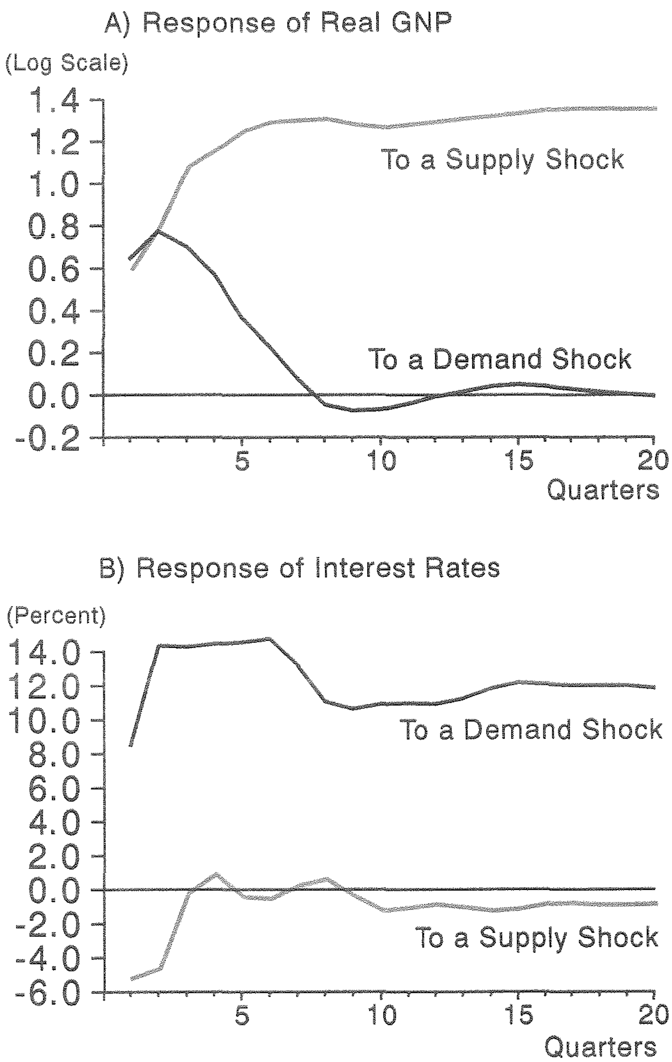
Econometric Method

An alternative approach to direct measurement is to estimate econometrically the demand and supply shocks that have influenced the aggregate macroeconomic time series data. The method used in this paper is that of Blanchard and Quah (1989) and involves estimating a vector autoregression under the assumptions that supply shocks can have long-run effects on real variables, while demand shocks can have only temporary effects.

Using these assumptions, it is possible to obtain estimates of demand and supply shocks from a VAR containing two variables: the rate of growth of per capita real GNP, y, and the quarterly change in the three-month Treasury bill rate, i.[14] Two types of (unobserved) structural disturbances are assumed to affect these variables. We identify these disturbances, respectively, with aggregate demand and aggregate supply. This procedure is described in the Box.

As noted earlier, the purpose of obtaining estimates of demand and supply shocks is to see if they help explain the dynamic relationship between inflation and unemployment. We chose to estimate the shocks independently of the inflation and unemployment data. Our concern was that if we estimated the shocks within the context of a model of inflation and unemployment, we would run an unacceptably high risk of capturing spurious correlations between the shocks, on the one hand, and inflation and unemployment, on the other hand. Of course, we recognize that output and interest rates are related to inflation and unemployment. However, by not using the latter two variables directly in the estimation of the shocks, we have reduced the chance of obtaining spurious results.

Chart 2
Dynamic Responses



Properties of the Estimated Shocks

To provide an indication of the nature of the shocks we estimated, Chart 2 presents plots of the dynamic effects of the typical demand and supply shocks on real GNP and the nominal interest rate over the 1955–89 period. Real GNP increases almost monotonically in response to a positive supply shock, growing rapidly in the first year following the shock, and then slowing down to its new long-run level about four years after the initial shock. Although the steady-state growth rate of real GNP is not affected by the supply shock, the *level* of real GNP remains permanently higher in steady-state.

The response of real GNP to demand shocks peaks in about two quarters and then dies out after approximately two years. In the case of a demand shock, by construction, both the level and the growth rate of real GNP are left unchanged in steady state.

In contrast, as shown in the second panel of Chart 2, a positive demand shock permanently raises the nominal interest rate. Supply shocks appear to have little or no effect on the nominal interest rate.

Table 1
Variance Decompositions
(Percent)

Forecast Horizon	Real GNP		Interest Rate	
Quarters Ahead	Supply Shocks	Demand Shocks	Supply Shocks	Demand Shocks
1	46	54	27	73
2	49	51	15	85
4	65	35	7	93
8	83	17	3	97
12	89	11	2	98
60	98	2	1	99

Estimating Demand and Supply Shocks From A Vector Autoregression

We employ moving-average representations of the quarterly changes in real per capita GNP (y) and the three-month Treasury bill rate (i) to illustrate how estimates of the underlying disturbances (v and e) are obtained. For illustrative purposes, we introduce dynamics into a simplified version of our model by including only one lag of v in each equation, although the estimated model contains more lags.

Assume, then, that the behavior of y and i is described by:

$$y_t = a_1 e_t + b_1 v_t + c_1 v_{t-1} \quad \text{(B1)}$$

$$i_t = a_2 e_t + b_2 v_t + c_2 v_{t-1} \quad \text{(B2)}$$

In order to study the dynamics of this system, estimates of the coefficients in equations (B1) and (B2) are needed. By placing certain identifying restrictions on the unobserved disturbances e and v, we are able to obtain estimates of the impact of each of them from observations on y and i. In traditional fashion, we assume that e and v are uncorrelated with each other and have unit variance. In addition, e and v are assumed to be serially uncorrelated. Given the representations in equations (B1) and (B2), these assumptions imply the following identifying restrictions:

$$\sigma_y^2 = a_1^2 + b_1^2 + c_1^2$$

$$\sigma_i^2 = a_2^2 + b_2^2 + c_2^2$$

$$\sigma_{y_t, i_t} = a_1 a_2 + b_1 b_2 + c_1 c_2$$

$$\sigma_{y_t, i_{t-1}} = c_1 b_2$$

$$\sigma_{y_{t-1}, i_t} = b_1 c_2$$

where σ_y^2 is the (observed) variance of y, σ_i^2 is the variance of i, $\sigma_{y,i}$ is the contemporaneous covariance of i and y, and the other variances are defined similarly.

So far, there are five conditions and six coefficients to estimate. One more restriction is needed to identify the model. Our final restriction comes from the assumption that v has no long-run effect on output; that is,

$$b_1 + c_1 = 0. \quad \text{(B3)}$$

This restriction, together with the conventional restrictions on variances and covariances, is just sufficient to identify the unobserved shocks from observations on y and i. The restriction defined by this final assumption is consistent with our interpretation of the underlying structural disturbances: v can be interpreted as an aggregate-demand shock, since it can have no long-run effect on y. In other words, the permanent level of real GNP is determined by real factors and can be only temporarily disturbed by aggregate demand factors. The effects of aggregate supply shocks will be captured by e, since these shocks are permitted to have permanent effects on the real variable, y.

In practice, these restrictions can be imposed in one of two ways. The first way—followed by Blanchard and Quah (1989)—is to estimate the usual vector autoregression (which involves regressing both y and i on lagged values of both variables) to obtain the variance-covariance matrix of the residuals. This matrix then is transformed to satisfy the restrictions given above. The second method (employed here) is due to Shapiro and Watson (1988), and is easier to implement. The two estimated equations are specified as

$$y_t = \sum_{j=1}^{n} \beta_{1,j} y_{t-j} + \sum_{j=0}^{n-1} \beta_{2,j} \Delta i_{t-j} + \epsilon_{1,t} \quad \text{(B4)}$$

$$i_t = \sum_{j=1}^{n} \gamma_{1,j} y_{t-j} + \sum_{j=1}^{n} \gamma_{2,j} i_{t-j} + \gamma_3 \epsilon_{1,t} + \epsilon_{2,t} \quad \text{(B5)}$$

As explained in detail in Shapiro and Watson, the inclusion of the interest rate variable in second-difference form[1] in the output equation is key to isolating a series that has only temporary effects on output. This restriction corresponds to (B3) above.[2]

1. Inclusion of second differences is equivalent to including the first differences of the interest rate with the restriction that the coefficients sum to zero.

2. Two technical points are worth mentioning. First, inclusion of the contemporaneous interest rate term in the output equation implies that we must employ instrumental variables techniques in estimating the equation. We used lagged values of the variables in the model as instruments. Second, once the interest rate terms are included in this way, a Choleski decomposition can be used to recover ϵ_1 and ϵ_2. (See Shapiro and Watson for details.)

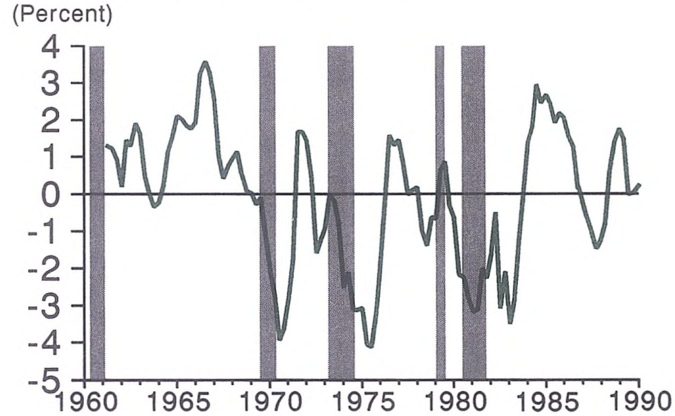
Table 1 presents the associated variance decompositions, which show the relative importance of demand and supply shocks in explaining the unpredictable movements in real GNP and the nominal interest rate. Demand shocks account for slightly more than one half of the unpredictable variation in output one quarter ahead, and about one third of that variation four quarters ahead. The remainder of the variation is accounted for by supply shocks. Thus, both demand and supply shocks play important roles in causing short-run fluctuations in real output. However, by construction, the long-run movements in the level of real GNP are the result only of supply shocks.

With respect to movements in interest rates, in contrast, demand shocks are much more important, accounting for about three fourths of the unpredictable interest rate variation one quarter ahead, and nearly all the variation at horizons of one year and beyond.

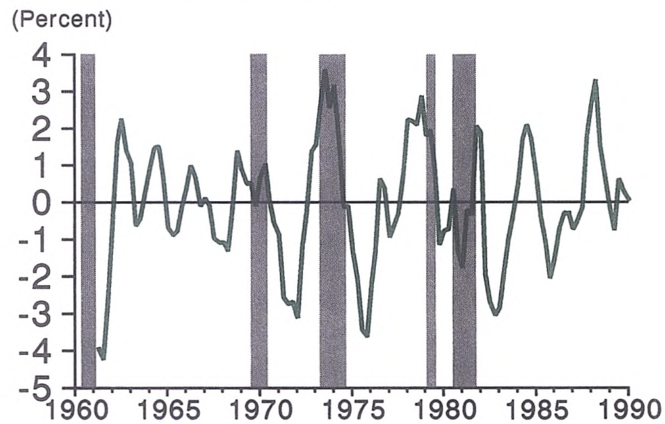
Chart 3 shows the estimated quarter-by-quarter effects of demand and supply shocks on real GNP over the period from 1960:Q1 to 1989:Q3.[15] These effects seem broadly consistent with the conventional interpretation of the major events in the period covered. As shown in the top panel, supply shocks were responsible for much of the above average economic growth during the 1960s, coinciding with the well-known productivity surge in that period. They also played a role in the 1973–75 and 1980–82 recessions, and may therefore be associated with the large oil shocks in those periods. Consistent with the recent history of monetary policy, the estimates shown in the bottom panel suggest that contractionary aggregate demand shocks played substantial roles in both the 1973–75 and the 1980–82 recessions, and that an expansionary demand shock was important in the late 1970s, when inflation accelerated.

Chart 3
Historical Decomposition of Real GNP Growth

A) Effect of Supply Shocks



B) Effect of Demand Shocks



III. Explaining Unemployment and Inflation

Having obtained measures of the supply and demand shocks during the 1960–89 period, our objective is to assess the relative importance of each of these shocks in explaining movements in inflation and unemployment. For this purpose, we estimated separate equations for unemployment and inflation as functions of current and lagged values of the estimated demand and supply shocks.

The equations are:

$$\pi_t = \sum_{i=0}^{16} \alpha_i s_{t-i} + \sum_{i=0}^{8} \beta_i d_{t-i} + \theta \bar{m}_t \qquad (1)$$

$$u_t = \sum_{i=0}^{4} \gamma_i s_{t-i} + \sum_{i=0}^{8} \delta_i d_{t-i} + \lambda u_{t-1} \qquad (2)$$

where π_t and u_t denote the rates of inflation and unemployment, respectively, and d_t and s_t are the (zero mean) demand and supply shocks. The inflation equation contains 16 lags of the supply shock variable and eight lags of the demand shock variable. These lag lengths were selected by doing F-tests on the relevant variables, four lags at a time. $\bar{m}_t$ denotes the average rate of M2 growth over the prior five years. It is included to allow the trend of inflation to move over the sample.[16] The demand and supply shock terms, then, explain deviations of inflation from the trend rate. Inclusion of $\bar{m}$ reduces the available sample size (compared with the sample used for the VAR above) because data for M2 begin only in 1959. Taking into consideration the lags in the model, 1965 is the earliest date at which we could begin the sample for the inflation equation.

The unemployment equation contains four lags of the supply shock, eight of the demand shock, and a lagged dependent variable. Without the latter variable, the significant lags on supply shocks in the unemployment equation were extremely long—at least 50 quarters. Thus, the lagged dependent variable was used to save degrees of freedom.[17]

In addition, we estimated

$$m_t = \sum_{i=0}^{8} \psi_i s_{t-i} + \sum_{i=0}^{4} \phi_i d_{t-i} + \sum_{i=1}^{3} \rho_i m_{t-i} \qquad (3)$$

where m_t denotes the rate of growth of M2. The M2 equation is required for dynamic simulations of the inflation equation, since the latter contains the 5-year average growth rate of M2 as a regressor. Equation 3 also contains constant dummy variables to eliminate the following observations from our sample: 1980:Q2 and 1980:Q3, because of the imposition and removal, respectively, of the Carter credit control program; and 1982:Q4 and 1983:Q1, because of the introduction of MMDAs. Finally, we also allowed the intercept term of the equation to change following the introduction of MMDAs.

Table 2

Summary Statistics From Regressions

Statistics	Dependent Variable: Unemployment Rate	Inflation	M2 Growth
Marginal Significance Levels of:			
F_1	.01	.01	.01
F_2	.01	.01	.01
t_1	—	.01	—
Adj. R^2	.99	.67	.79
SEE	.18	.40	.40
AR(1)	—	.41	—
(t-statistic)	—	(3.8)	—

F_1 is F statistic for null hypothesis that supply shocks have no impact on relevant variable.

F_2 is F statistic for null hypothesis that demand shocks have no impact on relevant variable.

t_1 is t-statistic for null hypothesis that M2 growth has no impact on inflation.

Table 2 presents summary statistics on these equations. Both the estimated demand and supply shocks are highly significant in all three equations. The errors from the ordinary-least-squares estimates of the inflation equation show evidence of first-order serial correlation, and we applied a correction for this. The first-order autocorrelation coefficient (AR(1)) estimate of 0.41 in the inflation equation compares to the AR(1) estimate of 0.75 in the raw inflation data, so our explanatory variables account for some, but not all, of the serial correlation in inflation.

To provide a better idea of the fit of these specifications, Chart 4 shows the actual values of inflation and the unemployment rate as well as dynamic simulations from our estimated equations. The equations do a good job of capturing the major swings in unemployment and inflation.

Chart 4
Dynamic Simulations

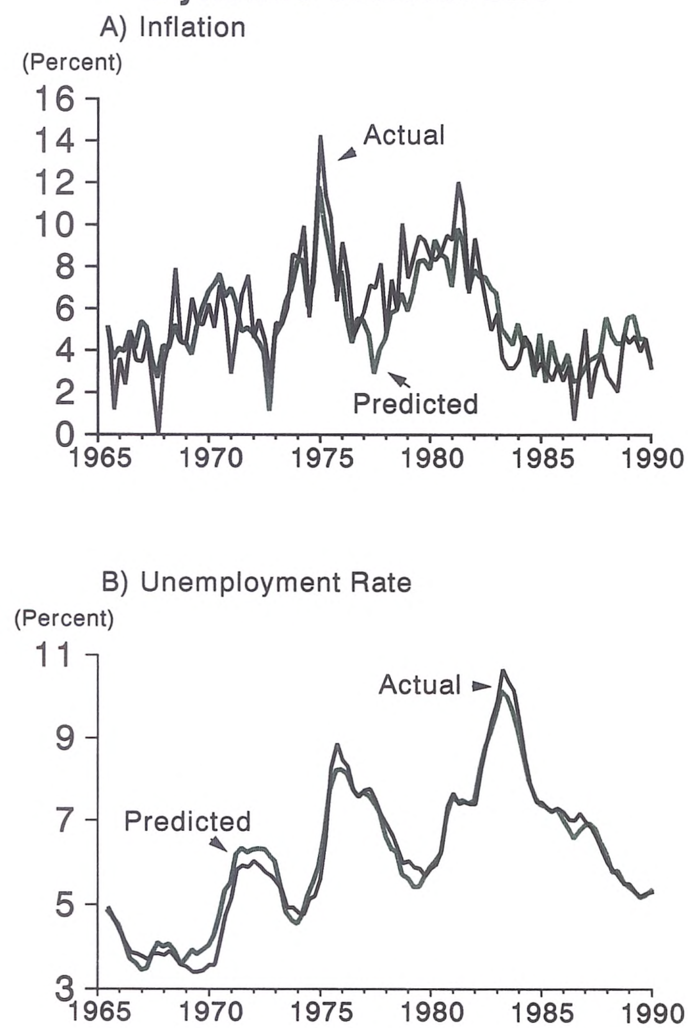


Chart 5
Dynamic Responses

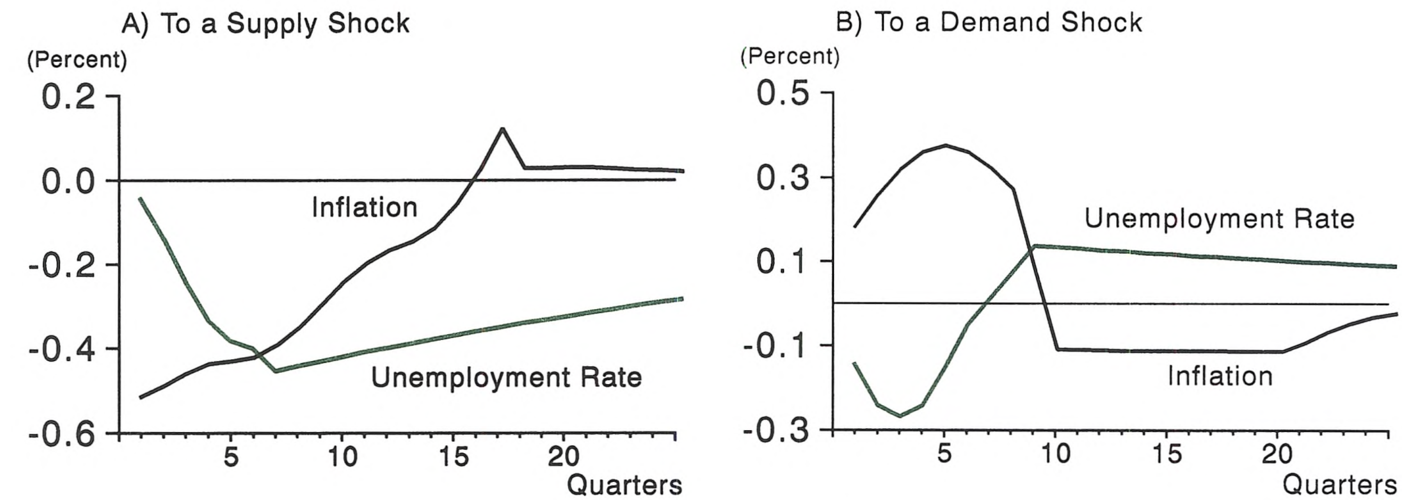


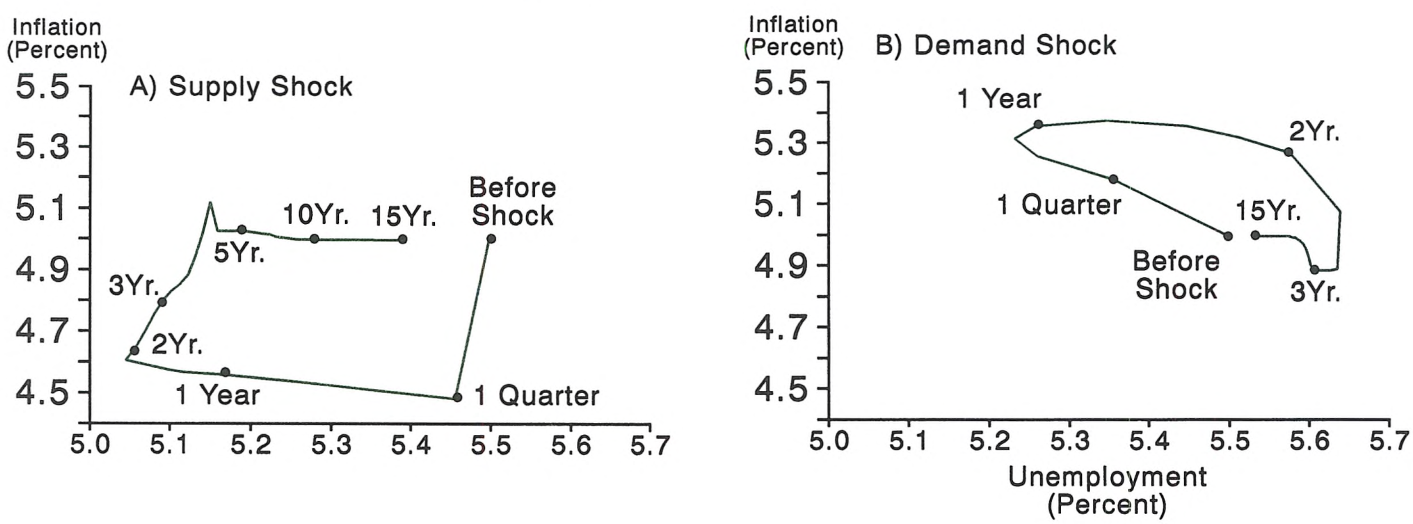
Chart 5 presents the estimated responses of inflation and unemployment to (one-standard-deviation) positive demand and supply shocks.[18] These dynamic responses have the signs predicted by theory. A positive demand shock reduces unemployment and raises inflation, while a positive supply shock reduces both unemployment and inflation.

Clockwise Loops

Chart 6 plots these dynamic responses in inflation-unemployment space. For illustrative purposes, we assume that the unemployment rate initially is 5.5 percent and the rate of inflation is 5.0 percent. The left panel shows the effects of a positive supply shock. The immediate response is a reduction in the inflation rate, after which unemployment gradually declines. The inflation rate moves back to its original level in two to four years after the shock, but the unemployment rate takes much longer to get back to its original level. Thus, even supply shocks lead to clockwise loops.

As shown in the right-hand panel, a demand shock initially has a larger impact on the unemployment rate. Unemployment reaches its minimum in less than a year, but by the end of the second year has risen back to its original level. The inflation rate rises as the unemployment

Chart 6
Dynamic Effects of Shocks on Unemployment and Inflation



rate declines, and remains high for nearly a year after the unemployment rate has returned to its original level. Thus, demand shocks produce the temporary trade-off predicted by the Phillips curve. Note that the effects of supply shocks evolve more slowly and take longer to be completed than those generated by demand shocks.

The loops in Chart 6 demonstrate why it is not possible to develop simple rules of thumb to judge future inflation based upon current observations of the level of unemployment. Any given rate of unemployment could be followed by almost any rate of inflation depending on the source of the shock. Further, since the loops ultimately go back to their starting points, a particular rate of unemployment will be associated with different rates of inflation at different points in time.

For similar reasons, *changes* in the unemployment rate are unlikely to provide accurate information about future changes in inflation. Consider, for instance, the left-hand panel of Chart 6. A falling rate of unemployment may be followed either by rising inflation (as inflation moves back to its original level between the second and fourth years after the shock), or by no change in inflation (as unemployment gradually adjusts back to its original level after the fourth year). Thus, Chart 6 provides an illustration of the general principle that using one endogenous variable to draw inferences about another endogenous variable can be a tricky enterprise.

Simulating Inflation-Unemployment Loops

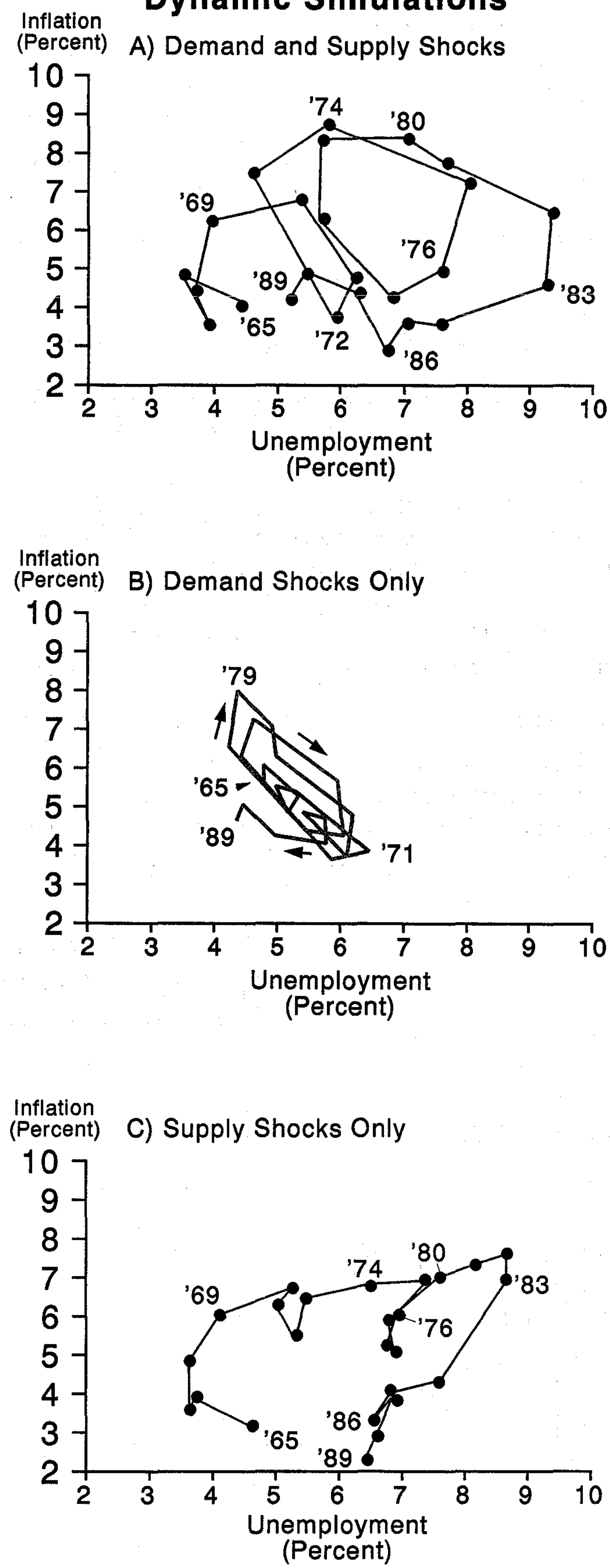
Chart 7 presents dynamic simulations of unemployment and inflation over the 1965–89 period, using the historical values of the estimated demand and supply shocks. The first panel shows the effects of both kinds of shocks over this period. The shape of our simulated loops is quite close to the actual data shown in Chart 1.

The second panel shows how unemployment and inflation would have evolved if there had been no supply shocks over this period. As expected, we obtain negatively sloped loops, with the number of loops attesting to the relatively short period over which the effects of a demand shock dissipate.

The third panel shows what would have happened if there had been no demand shocks over this period. The plot shows little tendency to loop around and come back to its original position, reflecting the long-lived effects of supply shocks.

Supply shocks have moved the unemployment rate and inflation over a much wider range than have demand shocks. Thus, they account for more long-run volatility in

Chart 7
Dynamic Simulations



these variables. Supply shocks are estimated to have caused inflation and unemployment to move within ranges that are 5.3 and 5.1 percentage points wide, respectively. The comparable figures for demand shocks are 4.3 and 2.2 percentage points.

In Section II, we discussed how our decomposition of movements in output into those caused by demand and supply shocks compared with conventional wisdom regarding the events over our sample period. Using Chart 7, we can now repeat this exercise in terms of combinations of inflation and unemployment. The largest movements that are estimated to have been caused by demand shocks occurred in 1977–79, and in 1980–83. In the earlier period, widely recognized as one of excessively expansionary monetary policy, the estimates suggest that demand shocks raised inflation by about 3½ percentage points and lowered the unemployment rate by about one percentage point. In the latter period, the Federal Reserve adopted reserves-oriented monetary policy procedures to reduce inflation. We estimate that this negative demand shock reduced inflation by about 4½ percentage points and raised the unemployment rate by nearly two percentage points in this period.

The largest supply shocks occurred in 1973–74, 1979–80, and 1983–86. The positive shock in the 1960s, presumably related to the large persistent productivity surge in those years, is mostly excluded by our 1965–89 sample period.[19] The 1973–75 and 1979–81 periods are associated with well known oil price shocks. According to our estimates, negative supply shocks raised inflation by about 1½ and two percentage points in these two periods, respectively, and raised unemployment by 2½ and 1¼ percentage points.

A large positive supply shock shows up in 1983–86. Any hypothesis concerning the source of this shock would be especially speculative. However, the period roughly corresponds with the cut in marginal tax rates in the early 1980s, which some have suggested was a supply-side source of rapid investment and increased work effort. In addition, rapid technological change in personal computing appears to have begun in the early 1980s, and this factor could be related to the estimated positive supply shock.

IV. Policy Issues

A major long-run goal of U.S. monetary policy is to eliminate inflation.[20] One way to attempt to meet this goal is to use (formal or informal) forecasts of future inflation to judge the appropriateness of the current stance of monetary policy. For example, given the current stance of policy, a forecast of inflation for any period in the future that exceeds the inflation goal would indicate that policy should be tightened. Our results suggest that the Phillips curve model of inflation could provide misleading signals under this forecast-oriented approach to policy.

The empirical importance of supply shocks as well as the long duration of their effects means that the unemployment rate can remain above or below its steady state value for long periods.[21] Consequently, to determine what a given rate of unemployment implies for future inflation, it is necessary first to determine the factors that are responsible for the prevailing unemployment rate. When analyzed in terms of the Phillips curve, supply-induced movements in the unemployment rate can lead to inappropriate policy actions. For example, a relatively low level of unemployment resulting from supply shocks offers little or no reason for concern about the potential for an acceleration of inflation. However, when viewed through the Phillips curve, such a change in the unemployment rate would suggest that policy should be tightened.

The preceding discussion is not meant to suggest that the Phillips curve model is inferior to other models of inflation that are currently available. On the contrary, the Phillips curve models appear to be at least as accurate at forecasting as the other available demand-side models. Stockton and Struckmeyer (1989), for example, support this conclusion with tests of forecasts from Phillips curve, monetarist, and monetary-misperceptions models. We have focused on the Phillips curve in this paper simply because it is incorporated into the large Keynesian-style "structural" models that are most widely used in macroeconomic forecasting.

Our major point is that there is good evidence that aggregate supply factors, in addition to aggregate demand factors, affect inflation dynamics in complex ways. Models that ignore part or all of these supply factors run the risk of making large errors in episodes when these supply shocks are important.

One response to this potential problem is to use an unrestricted vector autoregression for forecasting. VARs can capture both demand and supply factors, at least insofar as the average behavior of these shocks over the estimation period applies to the forecast period. Thus, this approach may provide more accurate forecasts on average; however, it appears susceptible to large errors in episodes involving large, atypical shocks.

Another response would be to develop a forecasting model along the lines of the approach used in this paper. Whether this approach would be fruitful is uncertain, since we are not aware that any such model has been built. In any event, given our finding that both demand and supply factors have been important in determining short- to intermediate-run macroeconomic developments over the past three decades, it would seem worthwhile to explore ways to disentangle the effects of these shocks in the context of forecasting future economic developments.

NOTES

1. Plosser (1989) questions the usefulness of distinguishing between demand and supply shocks, as well as the identification of real business cycle models with supply factors. Instead, he prefers to make a distinction between real and nominal factors.

2. For a discussion of the traditional Phillips curve, see Gordon (1982). Ball, Mankiw and Romer (1988) discuss the "new" Keynesian approach. Finally, for alternative theories concerning unemployment and inflation, see Lucas (1973) and Taylor (1980).

3. See Brayton and Mauskopf (1985) and Gordon (1982).

4. For analysis of the theoretical basis for the natural rate of unemployment, see Phelps (1970).

5. Within the context of a full Keynesian-style model, an oil shock can have an effect on the unemployment rate. At given nominal interest rates, for example, an adverse oil price shock could reduce real GNP by lowering business fixed investment and thus raise unemployment (via the IS and Okun's law relationships.) Note, however, that the increase in unemployment would feed into the Phillips curve like a demand shock: i.e., it would reduce inflation, tending to offset the direct upward pressure on prices from the oil shock.

6. For example, the inclusion of the relative price of oil occurred when the Phillips curve relationship became unstable in the mid 1970s following the oil embargo.

7. For discussions of real business cycle models and further references, see Plosser (1989) and Mankiw (1989).

8. Nelson and Plosser (1982).

9. See Huh (1990) and Cooley and Hansen (1989).

10. For unemployment to exist in equilibrium business cycle models, we need to allow for heterogeneity of firms or workers, necessitating job search. For a discussion, see Blanchard and Fischer (1989), pp. 346-350.

11. For a discussion of financial deregulation and its adverse effects on the stability of the monetary aggregates, see Simpson (1984). These developments do not imply, however, that there necessarily has been a change in the long-run relationship between M2 and inflation.

12. For discussion of issues in measuring the budget deficit and further references, see Gramlich (1989), Barro (1989), Bernheim (1989), and Eisner (1989).

13. Boschen and Mills (1988) have attempted to relate real shocks to various economic time series.

14. In an earlier paper, Judd and Trehan (1989), we used a five variable VAR to analyze unemployment rate dynamics. Using real GNP, the unemployment rate, a short-term nominal interest rate, the ratio of U.S. real exports to imports, and working-age U.S. population, we allowed for four different kinds of shocks—domestic technology, labor supply, (two different) demand shocks, and a foreign shock. That paper focused on relationships between these shocks and the unemployment rate, and did not explicitly analyze the inflation rate within the model.

15. These effects are obtained by multiplying the coefficients in the impulse response functions by the appropriate historical shocks as measured by the model. We use a forecast horizon of 40 quarters for this purpose, which moves the starting date of our sample to 1960:Q1.

16. As noted earlier, financial deregulation has made the relationship between M2 and inflation more susceptible to short-run disturbances. However, such disturbances can be expected to be internalized within the five-year-average observations used in equation (1).

17. Inclusion of the lagged dependent variable does not lead to demand shocks having long-lived effects on the unemployment rate because the later lags on the demand shocks have negative coefficients.

18. The estimated impulse response functions for inflation are noticeably jagged. Consequently, for the purposes of Charts 5 and 6, but not elsewhere in the paper, both the inflation and unemployment equations were re-estimated after imposing smoothness priors. For a discussion of these priors see Shiller (1973).

19. As discussed above, the inclusion of M2 forces us to shorten our sample period.

20. See Greenspan (1990) and Parry (1990).

21. When interpreted within the context of a Phillips curve equation, these supply-induced movements in the unemployment rate could appear to be changes in the so-called natural rate of unemployment.

REFERENCES

Ball, Lawrence, N. Gregory Mankiw, and David Romer. "The New Keynesian Economics and the Output-Inflation Trade-off," *Brookings Papers on Economic Activity*, 1988.1.

Barro, Robert J. "The Ricardian Approach to Budget Deficits," *The Journal of Economic Perspectives*, Volume 3, Number 2, pp. 37-54.

Bernheim, Douglas B. "A Neoclassical Perspective on Budget Deficits," *The Journal of Economic Perspectives*, Volume 3, Number 2, pp. 55-72.

Blanchard, Olivier J. and Stanley Fischer. *Lectures on Macroeconomics*. Cambridge, Mass.: The MIT Press. 1989.

Blanchard, Olivier J. and Stanley Fischer. *Lectures on Macroeconomics*. Cambridge, Mass.: The MIT Press. 1989, pp. 520-23.

Boschen, John F. and Leonard O. Mills. "Tests of the Relation Between Money and Output in the Real Business Cycle Model," *Journal of Monetary Economics*, 22, 1988.

Brayton, Flint and Eileen Mauskopf. "The Federal Reserve Board MPS Quarterly Econometric Model of the U.S. Economy," *Economic Modeling*, July 1985.

Cooley, T.F., and G.D. Hansen. "The Inflation Tax in a Real Business Cycle Model," *American Economic Review*, Summer 1989, pp. 733-48.

Eisner, Robert. "Budget Deficits: Rhetoric and Reality," *The Journal of Economic Perspectives*, Volume 3, Number 2, pp. 73-93.

Gordon, Robert J. "Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment," in M. Bailey, ed., *Workers, Jobs and Inflation*. Washington, D.C.: The Brookings Institution, 1982, pp. 89-158.

Gramlich, Edward M. "Budget Deficits and National Saving: Are Politicians Exogenous?," *The Journal of Economic Perspectives*, Volume 3, Number 2, pp. 23-35.

Greenspan, Alan. Testimony Before the Subcommittee on Domestic Monetary Policy, House Committee on Banking, Finance and Urban Affairs, U.S. House of Representatives, February 20, 1990.

Huh, Chan G. "Output, Money and Price Correlations in a Real Business Cycle Model," *Working Paper 90-02*, Federal Reserve Bank of San Francisco, May 1990.

Judd, John P. and Bharat Trehan. "Unemployment Rate Dynamics: Aggregate-Demand and -Supply Interactions," *Economic Review*, Federal Reserve Bank of San Francisco, Fall 1989, pp. 20-38.

Kydland, F.E. and E.C. Prescott. "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1982, pp. 1345-1370.

Lucas, Robert E. "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63, 1973.

Mankiw, N. Gregory. "Real Business Cycles: A New Keynesian Perspective," *The Journal of Economic Perspectives*, Volume 3, Number 3, Spring 1989, pp. 79-90.

Motley, Brian. "Has There Been a Change in the Natural Rate of Unemployment?," *Economic Review*, Federal Reserve Bank of San Francisco, Winter 1990, pp. 3-16.

Nelson, Charles R. and Charles I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 1982.

Parry, Robert T. "Price Level Stability," *FRBSF Weekly Letter*, Federal Reserve Bank of San Francisco, March 2, 1990.

Phelps, Edmund S. *Microeconomic Foundations of Employment and Inflation*. New York: W.W. Norton, 1970.

Plosser, Charles I. "Understanding Real Business Cycles," *The Journal of Economic Perspectives*, Volume 3, 1989, Number 3, pp. 51-77.

Shapiro, Matthew D. and Mark W. Watson. "Sources of Business Cycle Fluctuations," *NBER Macroeconomics Annual 1988*, Cambridge, Mass.: M.I.T. Press.

Shiller, Robert J. "A Distributed Lag Estimator Derived from Smoothness Priors," *Econometrica*, 1973, pp. 775-788.

Simpson, Thomas D. "Changes in the Financial System: Implications for Monetary Policy," *Brookings Papers on Economic Activity*, 1, 1984.

Stockton, David J. and Charles S. Struckmeyer. "Tests of the Specification and Predictive Accuracy of Nonnested Models of Inflation," *The Review of Economics and Statistics*, 1989, pp. 275-283.

Taylor, John B. "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, February 1980, pp. 1-23.

Imperfect Information and the Community Reinvestment Act

William C. Gruben, Jonathan A. Neuberger, and Ronald H. Schmidt

The authors are, respectively, Senior Economist and Policy Advisor, Federal Reserve Bank of Dallas; Economist, Federal Reserve Bank of San Francisco; and Senior Economist, Federal Reserve Bank of San Francisco. The editorial committee members were Reuven Glick, Chan Huh, and Elizabeth Laderman.

The Community Reinvestment Act has been used as a vehicle to increase lending to low-income neighborhoods. In this article, a conceptual framework is developed to evaluate the effect of CRA on bank portfolios. Results suggest that CRA will boost lending to low-income neighborhoods, with the cost of achieving the social goal of more even lending borne by bank customers and owners. The increase in lending to low-income neighborhoods is reinforced by an information effect: as banks expend greater effort searching for high-quality, low-income borrowers, this increased knowledge about the area reduces risks in that area.

Over the last two decades, community groups and Congress have expressed concern about "inadequate" lending to disadvantaged neighborhoods. These groups have argued that restrictive lending practices have led to decay and blight in some neighborhoods because qualified borrowers who would have made improvements could not get loans.

The arguments have centered around "redlining," a practice whereby a financial institution indiscriminately limits loans for the purchase of property in certain "undesirable" neighborhoods within its market area. According to this practice, lenders are alleged to deny loan applications for purchases in those geographic areas, regardless of the credit worthiness of the individual borrower. A number of statistically-based studies that purported to show evidence of redlining were influential in securing passage of the Community Reinvestment Act (CRA) as part of the Housing and Community Development Act of 1977.

Since 1977, redlining, or the absence of redlining, has occupied much of the attention of those investigating CRA-related issues. In practice, however, protests by community groups based on CRA grounds have delayed or prevented bank mergers and acquisitions even without specific proof that a particular bank has engaged in redlining. Moreover, the implementation of CRA has tended to encourage banks to make more loans in certain neighborhoods, regardless of whether a given bank has been shown to have engaged in redlining. Clearly, then, CRA has implications for banks beyond its anti-redlining provisions. Consequently, this article seeks to evaluate the effects of CRA within this broader context, leaving aside the questions whether CRA is necessary or whether redlining occurs. We develop a model to explain, first, why a bank might have different lending policies for different neighborhoods, and second, how the current approach to CRA's enforcement affects a bank's decisions in this regard.

Differential neighborhood lending patterns can be shown to be a rational response to an environment in which the costs of acquiring information are high and neighborhoods differ widely in their average default risks. The existence of CRA suggests that society views these differential lending patterns as "suboptimal," and CRA enforcement essentially requires financial institutions to bear the social cost of providing low-income borrowers greater access to funding. Specifically, banks are encouraged under CRA to increase lending in low-income neighborhoods. To do so, they may incur higher costs investigating the credit worthiness of potential borrowers in these areas than would be optimal from the perspective of profit maximization. CRA trades off bank profits for the social benefits derived from greater access by borrowers in disadvantaged neighborhoods.

This view of CRA, as a mechanism to induce banks to increase lending in low-income neighborhoods, is consistent with a variety of institutional responses to CRA. These responses can be interpreted as efforts to reduce the costs that kept banks from making loans prior to CRA. Among these responses are pooling agreements by banks and cooperative efforts between banks and community groups.

We present an imperfect information model in Section I to demonstrate conditions under which banks would choose to allocate credit across neighborhoods on the basis of average neighborhood characteristics. The effects of changes in information on the allocation of funds across neighborhoods are explored in Section II. The role of CRA in the lending decision is highlighted in Section III, where it is introduced as an additional constraint on the bank's choice. We discuss institutional arrangements that have emerged to minimize the cost burden of CRA in Section IV, and draw conclusions in Section V.

I. An Imperfect Information Model

The operations of a bank are based on a broad spectrum of factors and motivations, including maintaining the goodwill of customers, serving the needs of the community, and providing returns to investors. A central feature of commercial banks is their role as credit intermediaries. Banks develop expertise in evaluating the credit worthiness of borrowers and enjoy economies of scale in monitoring loans to ensure prudent behavior on the part of borrowers. The extent to which this monitoring activity is not performed easily or as efficiently by other participants in the credit markets determines the market share of banking relative to direct placement activity.[1]

Portfolio diversification is another factor in bank lending decisions. Every loan faces some default risk associated with the particular characteristics of the borrower or the project as well as overall economic conditions. A bank can reduce these risks, however, by diversifying its portfolio since the specific risks of every project are not perfectly correlated, and these risks tend to offset one another. As a result, the total risk of a diversified portfolio is generally less than that of an undiversified one.

Diversification can occur along many different dimensions, such as across industries (agriculture, manufacturing, services, etc.) and size categories (large corporations or single proprietorships), and across general classifications of customers (residential, industrial, commercial, etc.). Diversification also can be accomplished geographically, by lending to similar customers in different markets or neighborhoods.

As commonly expressed, the problem of "socially inadequate" lending to particular neighborhoods may arise because the costs of identifying good loans in certain areas outweigh the advantages that would be gained through greater geographic diversification. This form of credit rationing, sometimes referred to as "rational redlining,"[2] occurs when banks restrict lending or are less aggressive in marketing loan products in certain neighborhoods because the costs of identifying the qualified loans are too high to be profitable. Thus, the lack of readily available, complete information can affect the allocation of credit across neighborhoods.

In this model, we assume a bank operates in a geographic lending market comprising two neighborhoods, which we denote by the subscripts R (for "rich") and P (for "poor"). The bank divides its portfolio of loans between the two neighborhoods, with proportion θ allocated to neighborhood P, and $(1-\theta)$ to neighborhood R.

We define i_P as the contractual interest rate on neighborhood P loans and i_R as the contractual interest rate on loans to residents of neighborhood R.[3] We assume that loan markets are competitive and that interest rates on loans are determined in these markets. Loan interest rates may be *affected* by the underlying risks of projects in the two neighborhoods, although they may not correct for risk differentials exactly. We assume that the risk of a loan project in neighborhood P exceeds that in neighborhood R. Consequently, the loan rate in neighborhood P may be higher than that in neighborhood R.

The limit on banks' ability to tailor rates to reflect fully the differences in neighborhood risks may arise for a variety of reasons, including transaction costs or pressures from social or regulatory groups.[4] Moreover, problems of adverse selection, where higher rates may attract less credit worthy borrowers, can prevent banks from adjusting interest rates to compensate fully for risk.[5,6]

Finally, because much of the attention in the application of CRA is focused on home mortgage lending, the model reflects some of the characteristics of that market. In particular, while lenders can in principle charge differential rates across areas, in practice they tend to have only slight variations in mortgage terms at any given point in time. In large part, this leveling of rates and terms results from the desire of lenders to create homogeneous mortgage contracts that can be resold in secondary markets.

Because of these factors restricting interest rate differentials, we assume that individual banks are price takers and that rates are set exogenously. Thus, in this model, i_P and i_R are not viewed as explicit choice variables by the bank.[7]

In the absence of defaults, total dollar returns for the bank on neighborhood P loans are $i_P \theta L$ where L is the total dollar volume of loans in the bank's portfolio. Similarly, the income on neighborhood R loans is $i_R(1-\theta)L$. To simplify the analysis, we assume that the volume of loans is given. We define units such that $L = 1$ and thus eliminate it from the two expressions above.[8]

Banks are assumed to maximize an objective function that trades off risk and return. The bank's perception of a loan's riskiness depends, first, on the actual distribution of potential rates of return to that loan project. This distribution depends on the interaction between specific characteristics of the project and the realization of future random events. (This distribution likely will change when the interest rate changes.) Even with full information about current conditions, the bank still faces risks from future events.

Although the bank cannot observe this distribution and therefore derive a true measure of the loan's actual riskiness, it can estimate a project's riskiness by obtaining information about the details of the individual project, details that are observable at some cost to the bank. Thus, a bank's estimate of a project's riskiness depends on the amount of information gathered. Essentially, as the bank invests in more information, its ability to distinguish among borrowers and projects rises, allowing it to restrict its portfolio to the lowest risk projects seeking loans at the given contractual interest rate.

Risk can be expected to differ across neighborhoods. We assume that, for a given interest rate, the variance of returns to loans in neighborhood P is higher than in neighborhood R. This effect could be expected if income streams in the poor neighborhood were more volatile. Moreover, if banks operate less in low income areas, as is alleged by CRA advocates, they can be expected to have less information about neighborhood P.

The bank's estimate of risk to loans in the two neighborhoods can be written as:

$$\mathrm{VAR}(r_P) = \sigma_P^2 + \lambda_P(I_P) \quad (1)$$

$$\mathrm{VAR}(r_R) = \sigma_R^2 + \lambda_R(I_R) \quad (2)$$

where r_P and r_R are actual returns on loans that would be expected at the prevailing interest rates. σ_P^2 and σ_R^2 represent the variances of loan returns in neighborhoods P and R, respectively, that result from unobservable factors. We can think of these measures as the full-information minimum variances of returns of the portfolio the bank would choose if it had all the information available about possible projects. For a given interest rate, these components of the variance are assumed to be fixed. The λ terms in equations (1) and (2) are the result of the component of risk caused by imperfect information, with $\lambda_P > \lambda_R$.[9] These terms are dependent on the information the bank obtains about the two neighborhoods, I_P and I_R. Our assumption that information reduces this component of risk suggests that $\lambda_P' < 0$ and $\lambda_R' < 0$. We assume that λ_P and λ_R are unaffected by information about the other neighborhood. Thus, λ_P is independent of I_R and λ_R is independent of I_P.[10]

Banks, therefore, can reduce loan risk by acquiring more information about projects and borrowers in the two neighborhoods so that they can weed out the higher-risk projects. If information gathering were costless, they would seek information about both neighborhoods until $\lambda_i = 0$.

But information gathering is not costless. Banks typically must set up an on-site branch or loan origination office, conduct local surveys regarding the values of neighborhood properties, and solicit and evaluate loan applications from neighborhood residents—all functions that entail significant expenditures by the bank.

We characterize these information gathering costs by the average cost functions, $C_P(I_P)$ and $C_R(I_R)$. Total information costs are equal to $C_P(I_P)\theta$ for neighborhood P and $C_R(I_R)(1-\theta)$ for neighborhood R. Typically, there are large initial fixed costs associated with the investment in information (such as setting up a branch), which lead to declining average costs over some range. We assume, however, that the marginal cost of obtaining information

Imperfect Information vs. Credit Rationing

The model presented in this paper is explicitly designed to focus on factors other than differential interest rates that can affect neighborhood lending patterns. The model stresses the role of imperfect information and the interaction of CRA restrictions.

Another model that offers complementary insights into the issue is the credit rationing model, of which work by Stiglitz and Weiss (1981) is often taken as a major point of departure. In that framework, it is possible to show why banks might not choose simply to compensate for higher risks in different neighborhoods by charging higher rates in those areas.

Credit rationing is said to exist when lenders refuse to grant credit to a borrower even when that borrower is willing to pay a higher interest rate. Such refusal is predicted when the lender does not expect to obtain its required return at any interest rate. In this context, it is important to understand the difference between the lender's expected rate of return and the interest rate charged on a loan. The expected rate of return is the interest rate charged net of the expected rate of default. At any given interest rate, an increase in the expected default rate means a decrease in the expected return on the loan.

This difference has important links to two ideas that are closely tied to theories of redlining: (1) information asymmetry and (2) adverse selection.

In cases of information asymmetry, lenders and borrowers have access to different sets of information regarding the credit quality of borrowers. While lenders may know that certain borrower characteristics are related to high default rates, they do not know all the characteristics of individual borrowers because borrowers can withhold information. A borrower who has characteristics that lenders associate with high credit risk has an incentive to hide this information. As a result, it is difficult for lenders to distinguish between a loan applicant who will default and one who will not. That is, borrowers who are observationally indistinguishable may be different in fact.

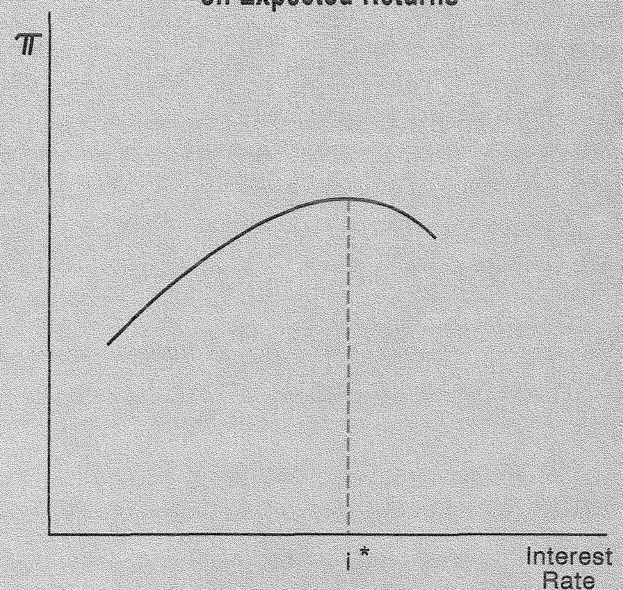
A related concept is the theory of adverse selection. According to this theory, different interest rates imply different pools of loan applicants. A rise in interest rates can induce an adverse change in the mix of applicants. Safe potential borrowers drop out of the market. These borrowers are discouraged by higher interest rates because these rates mean higher payments relative to the value of the loan project. On the other hand, borrowers who know they are unsafe are less concerned about the higher loan payments. They know they have a higher likelihood of default and, thus, a higher probability of not repaying the loan anyway.

Even though individual borrowers may have risk-related characteristics that lenders cannot identify, lenders understand the relationship between interest rates and the potential for adverse selection of borrowers. Lenders realize that beyond a certain point, increasing the interest rate on loans can mean an increase in defaults and thus lower expected returns. Therefore, lenders will not lend at rates above a certain maximum because the expected rate of return on lending falls.

This notion is depicted in Figure B1. The vertical axis represents the bank's expected return on loans (net of losses as a result of defaults), a concept analogous to the variable π in our model. The horizontal axis represents the contractual interest rate on loans. The expected return to the bank rises with the quoted interest rate until it reaches i^*. Below this interest rate, the greater interest income resulting from higher contracted interest rates dominates any increase in default rates resulting from the change in the quality of the pool of applicants. Above i^*, however, so many safe borrowers are discouraged from applying for loans that the expected return to the bank actually falls. At these high rates, there is a larger share of risky borrowers in the pool of applicants, and default rates can rise by more than the increase in interest income to the bank.

Figure B1

Effect of Adverse Selection on Expected Returns



In practice, information asymmetries and adverse selection confront lenders simultaneously. Suppose, for example, that lenders can distinguish between two groups of borrowers, a high-risk group and a low-risk group. Information asymmetries may make it difficult for the lender to differentiate among the members of each group (as in the previous discussion). But there generally is some set of observable characteristics of loan applicants that allows the lender to place some applicants in the low-risk group and the rest in the high-risk group. Suppose, in addition, that lenders know as much about members of the high-risk group as they know about members of the low-risk group.

Figure B2 shows how the existence of observably different groups of applicants can lead to credit rationing. The axes in Figure B2 are the same as in Figure B1. In this figure, the lender must derive a rate of return that is at least as high as its cost of funds, r_d (the rate the bank pays on deposits). At any given interest rate, borrowers in the low-risk group (Group I) provide a higher expected return to the bank than borrowers in the high-risk group (Group II). Moreover, the expected default rate for Group II is so high that no contracted interest rate yields an expected return sufficient to cover the lender's cost of funds. That is, all points on the curve that characterizes Group II are below r_d, the lender's cost of funds. In this case, members of Group II will be rationed out of the credit market. In the sense described by Jaffee and Stiglitz (1990), Group II borrowers are redlined, even though some Group II projects might yield sufficient returns to justify allocating credit.

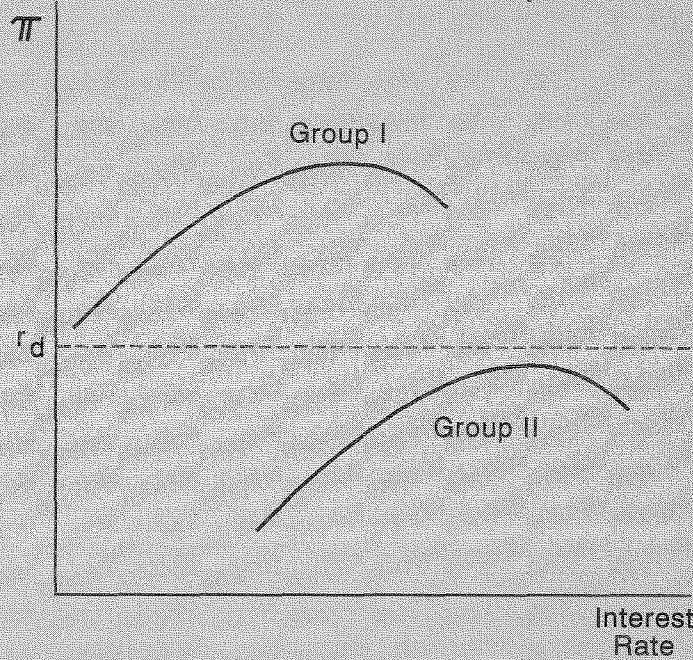
Credit Rationing vs. the Imperfect Information Model

The model in Section I is complementary to the standard credit rationing model. Stiglitz and Weiss (1987) focus on the optimal loan rates to charge different groups and the decision to allocate credit to those groups. In doing so, they assume that the bank cannot affect the expected return/loan rate trade-off through additional information gathering. In terms of our notation, their model assumes that $\lambda(I_i) = 0$.

In contrast, the model presented in Section I examines how banks can affect the expected return/loan rate trade-off through information gathering. To simplify the analysis, we ignore the factors that determine loan rates (the problem investigated by Stiglitz and Weiss). We thus do not provide a simultaneous determination of loan rates, portfolio shares, and information acquisition. We choose instead to focus on the influence of variance-reducing information on the credit allocation decision of the bank, taking contractual interest rates as given. This framework, we believe, provides a useful analytical tool to investigate how CRA affects bank lending decisions.

Figure B2

Credit Rationing with Different Borrower Groups



The presence of obtainable information alters the motivation for the differential treatment of borrowers in our model relative to the typical credit rationing story. In the credit rationing model, there is a fundamental difference in the default risks of the two groups of borrowers. Since one group of borrowers is riskier than another, lenders respond to this difference in default characteristics by curtailing credit to the riskier group. In our framework, the two neighborhoods differ in terms of *perceived* risks of default, with the difference in perceived risks exacerbated by information problems. In the absence of any information investment, the perceived risk in the rich neighborhood is smaller than the perceived risk in the poor neighborhood. Thus, as in the credit rationing model, the rich neighborhood receives more credit. However, since the bank can purchase information that reduces default risk, the cost of information affects the allocation of credit between the two neighborhoods.

To summarize, the model that relies on costly information offers one explanation for differential lending behavior across neighborhoods. When the perceived variance of returns in one neighborhood is significantly larger than in the other, the bank allocates less credit to the area with the higher risk. Thus, apparent redlining simply may be a rational response of banks to the problems raised by imperfect, and costly, information.

eventually rises as the desired type or quality of information becomes more difficult to obtain or evaluate. This produces U-shaped average and marginal cost curves.

We abstract from the deposit-taking activities of the bank, even though this activity may affect the bank's cost of funds and, thus, its profits. In effect, we assume a perfectly elastic supply of deposits at a risk-free rate of return to depositors, r_d.[11] As a result, we separate the bank's deposit-taking function from the business of making loans.

With the components described above, it is possible to state the bank's objective function as

$$\text{Max } \pi = i_P\theta + i_R(1-\theta) - \beta \text{ var}[r_P\theta + r_R(1-\theta)] - C_P(I_P)\theta - C_R(I_R)(1-\theta) - r_d \quad (3)$$

where π is the bank's adjusted return. π must be positive for the bank to operate. The bank chooses values of θ, I_P, and I_R that maximize this expression.

The objective function in (3) asserts that the bank seeks a balance between a portfolio's interest income, information costs, and variance. In this case, we assume that the adjusted return depends negatively on portfolio variance. This effect could arise for a variety of reasons, including risk aversion on the part of banks.[12] In addition, bankruptcy costs resulting from failed projects can be expected to make portfolio risk a negative factor to the bank. The third term, therefore, reflects a reduction in the bank's income from expected loan losses to its portfolio, which for simplicity is assumed to be a constant multiple, β, of the bank's portfolio variance.[13]

This formulation differs from other models of credit allocation. For example, as discussed in the box, "Imperfect Information vs. Credit Rationing," the standard credit rationing model focuses on the effects of asymmetric information on the determination of loan rates and the decision to exclude or ration credit to various borrower groups. The current work, which should be viewed as complementary to this credit rationing model, focuses on the process by which information gathering changes credit allotments. Moreover, the structure of this model is designed explicitly to model the effect of CRA, which is difficult to incorporate directly into the credit rationing framework. Nevertheless, many of the implications of the credit rationing model can be expected to carry over into this analysis as well.

The solution of (3) yields the following optimal conditions for a bank's portfolio allocation and information gathering:

$$\theta^* = \frac{i_P - i_R + C_R - C_P + 2\beta[\text{var}(r_R)]}{2\beta[\text{var}(r_P) + \text{var}(r_R)]} \quad (4)$$

$$C'_P(I^*_P) = -\beta\theta\lambda'_P(I^*_P) \quad (5)$$

$$C'_R(I^*_R) = -\beta(1-\theta)\lambda'_R(I^*_R) \quad (6)$$

Equations (4) to (6) represent equilibrium first-order conditions for the three choice variables, which are simultaneously determined and clearly interdependent. By totally differentiating equations (4) to (6), it is possible to solve for reduced form expressions that calculate the effect of changes in exogenous variables and model parameters on the equilibrium values of the choice variables. We present these comparative statics results in the Appendix.

The solution to the model suggests several factors that influence a bank's allocation of loans across neighborhoods:

- The bank lends a larger proportion of its loan portfolio in the poor neighborhood when the contractual interest rate on neighborhood *P* loans rises relative to that on neighborhood *R* loans.
- When the variance of neighborhood *P* returns falls relative to that in neighborhood *R*, the proportion of loans in neighborhood *P* rises. Assuming that the risk-adjusted return to lending in the rich neighborhood is higher before $\text{Var}(r_P)$ declines, the relative advantage of the wealthier neighborhood is eroded.[14] Similarly, factors that reduce the cost of obtaining information about neighborhood *P* relative to that for neighborhood *R* increase the proportion of loans to neighborhood *P*. Clearly, if information is less costly to obtain in one area, more information is acquired, thereby reducing the relative variance of returns to that neighborhood.
- The effect of an increase in risk aversion (or the expected default rate) is less clear and depends, among other things, on the spread between contractual interest rates and differences in variances of the two neighborhood returns. As β rises, the value of reducing the portfolio's variance rises, pushing the solution toward the minimum variance portfolio. For low initial levels of θ, the effect of an increase in β is to shift the portfolio toward neighborhood *P* loans to capture the advantages of portfolio diversification. At high values of θ, similar diversification incentives shift the loan portfolio toward neighborhood *R* loans.

The solution to the model is depicted graphically in Figure 1. The figure shows how different allocations across neighborhoods affect profits, holding constant the optimal quantities of information, I^*_P and I^*_R. The curve labelled π_P represents the portion of adjusted returns attributable to lending in neighborhood *P*. The π_R curve is the equivalent measure for neighborhood *R* loans. The curve marked π is the vertical sum of the π_P and π_R curves, representing the bank's total profits as a function of θ. At $\theta = 0$, the entire bank portfolio is allocated to neighborhood R loans and the total adjusted return is thus equal to π_R. Conversely, at $\theta = 1$, the bank's portfolio consists entirely of loans to neighborhood P projects and total returns are derived from π_P. The two functions, π_P and π_R are concave in θ.[15] Their curvature creates the total return function that, in the current figure, rises over some portion of values of θ, and then falls. The profit-maximizing bank chooses the highest point on the total return curve, with an optimal credit allocation equal to θ^*.

Factors that raise the marginal profit of neighborhood P loans relative to that of neighborhood R loans will increase θ^*. For example, an increase in i_P relative to i_R, or a decrease in the variance of r_P relative to that of r_R will rotate π_P upward, and tend to move θ^* to the right. An increase in β will increase the concavity of both profit functions, and θ^* will increase if it was very low initially and if the increased concavity raised the marginal profit of neighborhood P loans more than that for neighborhood R loans.

In the solution shown in Figure 1, neighborhood P is not redlined, that is, the optimal credit allocation implies $1>\theta>0$. It is possible, however, to derive a redlined solution in the current framework. Redlining will occur if the marginal profit from lending to the poor neighborhood, represented by the slope of the π_P function, is less than the absolute value of the slope of the π_R function at low levels of θ. In such a case, the total return function slopes downward over its entire length, with a maximum value occurring at $\theta = 0$. In this case, the profit-maximizing bank would allocate all of its loans to neighborhood R projects and would redline neighborhood P.

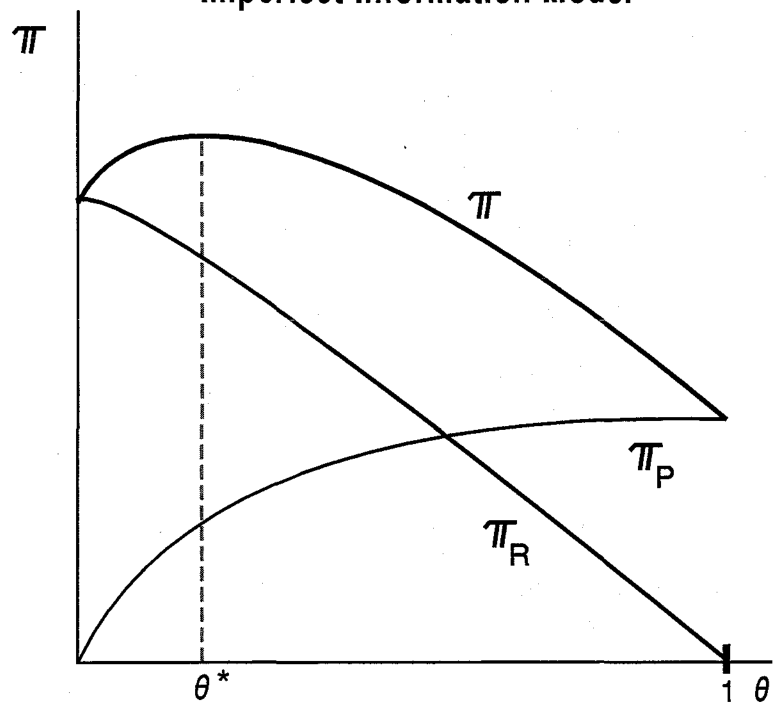
II. The Effect of Information

The equilibrium lending pattern across neighborhoods, shown in Figure 1, is directly affected by the information gathering process. As we demonstrate in this section, the equilibrium values of all three choice variables are highly interdependent, with optimal investment in information about both neighborhoods determined simultaneously with the decision about portfolio shares.

The optimal investment in information about the two neighborhoods is closely related to the portfolio allocation decision. The first-order conditions for I_P and I_R, equations

Figure 1

Optimal Credit Allocation in the Imperfect Information Model



(5) and (6), suggest that the bank should invest in information until the marginal cost of the last units acquired (the left-hand sides) just equals the value of the reduced portfolio risk (the right-hand sides). (Note in equations (5) and (6) that $\lambda_i' < 0$ and $C_i' > 0$).

Comparative statics results for the information choice variables, derived in the Appendix, show several factors that affect the optimal information acquisition:

• The amount of information acquired is positively related to the share of loans allocated to that neighborhood. Thus, θ and I_P move together. Whenever the bank allocates more of its loan portfolio to the poor neighborhood, it also is in the bank's interest to acquire more information about neighborhood *P* borrowers and projects. Because more of the portfolio is at risk in the neighborhood, the marginal benefit of information about that neighborhood rises. Any factor that raises the optimal value of θ, such as a relative increase in contractual interest rates on neighborhood *P* loans or a decrease in the risk of neighborhood *P* projects, also will induce the firm to obtain more information about the poor neighborhood.[16]

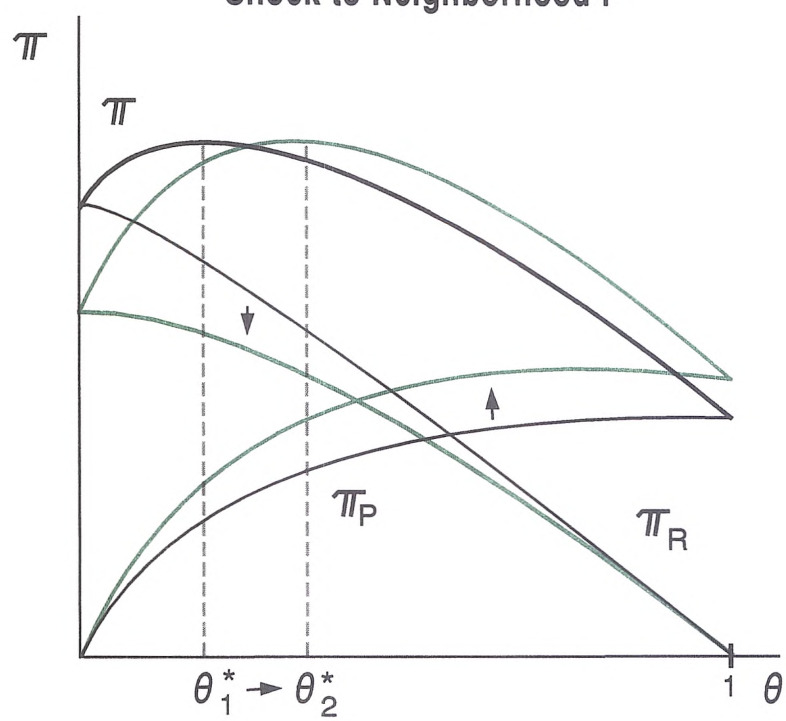
• The amount of information purchased also depends on the degree to which more information reduces portfolio variance. If additional information about neighborhood *P* becomes less valuable because some exogenous factor negatively affects the bank's perception about the neighborhood's risk, the marginal benefit falls. The bank then reduces the amount of information it obtains about the poor neighborhood, and allocates less of its portfolio to neighborhood *P* loans. The opposite effect occurs if the marginal benefit of I_R falls; that is, the bank responds by increasing its allocation of loans to the poor neighborhood and purchasing more information about neighborhood *P*. These results suggest that information investment in the two neighborhoods is a choice between substitutes: factors that raise the value of information in one neighborhood also reduce the relative value of information about the other neighborhood. This result is dependent on our assumption that information about one neighborhood does not affect the variance of projects in the other neighborhood.

• Factors that reduce the marginal cost of information for one neighborhood will increase investment in information in that neighborhood. As discussed in Section IV, pooling arrangements and collaboration with community groups can reduce the marginal cost of information to an individual bank. Banks will then increase information to balance marginal costs and benefits.

• Finally, the solution to the model depends on the relationship between defaults and portfolio variance. Obviously, if there were no defaults, (that is, $\beta = 0$) the bank

Figure 2

Effect of Positive Information Shock to Neighborhood P



would not invest in information at all. Information is only valuable for its variance-reducing content. As the sensitivity of adjusted returns to portfolio variance increases, that is, as β rises, the value of information rises, and the bank invests in more information about *both* neighborhoods.

The role of information gathering in our model can be seen in Figure 2. The black curves represent the initial equilibrium depicted in Figure 1. We suppose that a positive shock occurs to neighborhood P that raises the return to information. For example, a firm announces its intention to build a manufacturing plant in neighborhood P, with plans to hire many local workers. This announcement should lead to better income prospects for neighborhood P residents, and raise the credit quality of the pool of applicants in that neighborhood. The assumed shock to neighborhood P increases the bank's perception of the marginal benefits of additional information. (This implies that λ_P' increases in absolute value as a result of the announcement.)

The higher marginal benefit of information about neighborhood P induces more investment in that information. As a result, π_P shifts upward. The higher profit schedule, therefore, boosts the optimal θ. The rise in θ also boosts the marginal benefit of information in neighborhood P, while reducing the marginal benefit of information in neighborhood R, thereby reinforcing the initial information shock. The bank responds by increasing the information investment in the poor neighborhood even more, further shifting the π_P curve in Figure 2 up to the green line. At the same time, the marginal benefit to neighborhood R information falls relative to its marginal cost. The bank reduces its information investment in the rich neighborhood, resulting in a downward shift of the π_R curve in Figure 2. The new equilibrium implies that the bank responds to the positive shock in the poor neighborhood by raising the portion of its portfolio allocated to neighborhood P loans, increasing the amount of information purchased about neighborhood P, and reducing the information investment in neighborhood R.[17]

This imperfect information framework may shed light on empirical studies claiming to find evidence of neighborhood redlining. Given the cost of obtaining information, it would not be surprising to observe banks using general neighborhood characteristics to evaluate return functions, as well as to assess information costs. This type of information is widely available at relatively low cost. To the extent that racial and other social characteristics are correlated with overall economic variability, returns can appear to be a direct function of these characteristics. Moreover, as the costs of finding the lowest risk loans rise, the potential for credit rationing increases in neighborhoods with characteristics correlated with higher risks. Use of these characteristics as a screening device may therefore represent a first guess by lenders as to default risk in different neighborhoods. Once banks choose (or are induced) to make more substantial investments in information, however, the relevance of general neighborhood characteristics may give way to more costly borrower- or project-specific data that carry more information content. In the context of our model, CRA represents one such inducement.

III. The Role of CRA

CRA can be considered an additional regulatory constraint imposed on banks, thereby affecting their optimal portfolio allocation across neighborhoods. As we show in this section, the basic model presented in Section I can be easily modified to capture the essential features of CRA.

Congress passed the Community Reinvestment Act partially in response to community groups' claims that previous anti-discrimination laws had failed to keep banks from redlining. A financial institution is said to redline if it indiscriminately denies loans for the purchase of property in certain "undesirable" neighborhoods within its market area. In hearings prior to the drafting of what became the CRA, statistically-based studies were presented (New York Public Interest Research Group, 1977; National Peoples Action, 1976) that claimed to confirm the existence of redlining, despite the earlier passage of the Equal Credit Opportunity Act (1974) and the Home Mortgage Disclosure Act (1975).

It is true that those earlier laws may not provide effective sanctions against redlining. While the Equal Credit Opportunity Act prohibits discrimination in credit transactions on the basis of race, color, religion, national origin, sex, marital status, and age, it does not outlaw geographic discrimination. The Home Mortgage Disclosure Act requires financial institutions to disclose data on the volume of mortgage loans by census tract or zip code, but does not proscribe geographically discriminatory loan policies.

CRA requires federal regulators to motivate commercial banks and thrift institutions to meet community credit needs by considering a financial institution's record of community lending when they evaluate its applications for mergers or acquisitions. Members of the public also may formally protest an application if they think that the institution's record with regard to lending in certain neighborhoods is unsatisfactory.

The primary purpose of CRA, which is expressed in purposefully vague language, is subject to debate. An often-used, narrow interpretation is that CRA is an anti-redlining bill. In this view, the critical issue becomes one of identifying clear evidence that banks engage in irrational redlining. Otherwise, enforcement of CRA is not needed.

The evidence in redlining studies is inconclusive. Several studies have found average neighborhood racial and other social characteristics to be significantly correlated with lending activity even after controlling for a variety of other influences. However, these studies have been criticized for excluding important variables or using incorrect data—factors that can lead to overestimates of the importance of race or other social factors in neighborhood lending. (For a discussion of empirical studies of redlining, see the box entitled "Evidence of Redlining.")

CRA is viewed more broadly in this study, however. Rather than an anti-redlining law, CRA is viewed here as a mechanism to increase disadvantaged neighborhoods' access to credit whether or not redlining was actually occurring. This broader view of CRA is consistent with two recent developments. In February 1989, the Federal Reserve Board denied on CRA grounds an application by Continental Illinois to acquire another institution, even though Continental Illinois was not believed to be engaged in redlining *per se.*

Also, in 1989, the federal regulatory agencies[18] revised the guidelines for compliance with the CRA, and established more stringent and specific standards. However, because CRA requires that lending be consistent with "safety and soundness" considerations, even the new guidelines do not delineate an acceptable geographic pattern of lending. In the initial statement of the CRA, twelve criteria were to be used in evaluating a lender's record of compliance. One of the criteria states that regulators are to consider "the geographic distribution of the bank's credit extensions, credit applications, and credit denials." The 1989 amendment refers to "unwarranted geographic differences in lending patterns," and to "disparities in lending that do not appear to be attributable to safety and soundness considerations or to factors beyond an institution's control." What would make these geographic differences unwarranted or unattributable to safety and soundness considerations is not stated.

In light of this broader perspective on CRA, two alternative interpretations of the application of CRA have emerged. The regulatory mandate requires that financial institutions search harder for good loans in disadvantaged neighborhoods, but does not outlaw the rationing of credit or require banks to make riskier loans. The law imposes on banks and thrifts the costs associated with expending the effort to seek out high-quality borrowers in areas that are perceived as riskier.

In this "effort-oriented" approach, a bank is penalized for noncompliance with CRA if it demonstrates insufficient efforts to meet the credit needs of the community it serves. The penalties take the form of delays in processing, or even denial of, applications for mergers and acquisitions. CRA examination ratings consider the extent to which the bank conducts outreach programs, educates the public on its policies, and aggressively markets its products in low-income neighborhoods.

Effort, however, may not translate into a greater amount of funds lent to poorer neighborhoods. Recent CRA protests, therefore, have focused more on results than on effort. Challenges to bank mergers and acquisitions on CRA grounds have been raised when community groups have claimed that those institutions failed to meet a "socially acceptable" minimum level of lending in lower-income neighborhoods. In order to avoid the CRA "penalty," banks have responded with specific commitments of loan funds to those neighborhoods.

Modelling CRA

In modelling the effect of CRA, it is necessary to choose between the regulatory interpretation and the more recent results-oriented application of CRA. An "effort-oriented" approach would focus on the amount of information the bank acquires. In contrast, a "results-oriented" approach emphasizes θ, the proportion of the portfolio allocated to the low-income neighborhood. We have chosen to model the latter interpretation.

In the context of our imperfect information model, CRA has the effect of establishing a minimum proportion of the bank loan portfolio allocated to the poor neighborhood. The value of this minimum allotment is determined by a social welfare function that is exogenous to our model. We refer to the socially acceptable minimum level of credit allocated to the poor neighborhood as $\bar{\theta}$. We assume that CRA imposes a penalty on the bank (delays in processing applications, negative publicity, etc.) if it fails to allocate at least this proportion of its loan portfolio to neighborhood P loans. We characterize this penalty by the function:

$$d(\bar{\theta} - \theta) \qquad (7)$$

where $d(.) > 0$ when $\theta < \bar{\theta}$, and $d(.) = 0$ when $\theta \geq \bar{\theta}$.

Evidence of Redlining

Concerns over social responsibility in lending complicate the issue of redlining because the characteristics of a neighborhood in which a loan is made affect the likelihood of default. Barth, Cordes, and Yezer (1979) find that even after taking into consideration the effects of intercity differences in per capita income, foreclosure rates are higher in blighted neighborhoods, on properties in poor condition, and, among other things, on houses constructed of wood siding.

When financial institutions cannot vary mortgage rates among borrowers sufficiently to accommodate differences in the likelihood of default, it should not be surprising to see them lend in neighborhoods where the likelihood of default is low and avoid areas where default is more probable. That is, a competitive financial system can be expected to equalize expected rates of return on loans, after adjustment for probabilities of default. Thus, high denial rates in the high-risk neighborhood need not reflect redlining in the normally accepted sense of the term.

The empirical literature that attempts to identify redlining is subject to controversy. A common criticism (Benston, 1981; Sullivan and Pozdena, 1982; Jones and MacLennan, 1987) leveled at such research is that it often ignores factors that would motivate reasonable differences in mortgage lending activity among neighborhoods. That is, such literature is often said to give insufficient consideration to default-related factors such as those identified by Barth, Cordes, and Yezer (1979).

Moreover, reports that purport to confirm the existence of redlining and that were presented at the 1977 Hearings on Community Credit Needs that preceded enactment of the CRA are among those claimed (Benston, 1981) to be particularly susceptible to such criticisms.

In addition, a number of studies that do consider loan-risk factors as determinants of differentials in neighborhood lending patterns (Richardson and Gordon, 1979; Hutchinson, Ostas, and Reed, 1977; Ostas, Reed, and Hutchinson, 1979) fail to find evidence consistent with redlining as defined above.[1] In such studies, even though measures of mortgage extensions vary among neighborhoods, the variance can be explained chiefly by factors that would generate differential perceptions of loan risk among the neighborhoods.

Nevertheless, some recent research does point to the possible existence of redlining in the sense noted here. Bradbury, Case, and Dunham (1989) find that even after accounting for neighborhood differences in wealth, income, housing values, vacancy rates, the percentage of properties that are commercial and industrial, the presence of depository institution offices, and a number of other neighborhood characteristics, loan activity per number of Boston-area housing units is still negatively associated with the percentage of neighborhood resident population that is black.[2] Jones and MacLennan (1987) argue that they find evidence of redlining in Glasgow, Scotland, but they suggest that the process occurs through loan officer guidance of potential borrowers toward some neighborhoods and away from others. However, neither study identifies neighborhoods in which no mortgage lending took place, and a number of structure-specific default-related characteristics noted in other studies, such as Barth, Cordes, and Yezer (1979) are not considered.

In sum, redlining may or may not exist in a narrowly defined sense, but some geographic differentials in mortgage lending most certainly do arise. Moreover, the Community Reinvestment Act discourages such differentiation unless extremely compelling reasons dictate it, and imposes penalties when there is evidence to suggest that lending patterns differ significantly across geographic regions.

NOTES

1. Hutchinson, Ostas, and Reed (1977) note that their results "are consistent with the hypothesis that redlining takes place on the basis of risk aversion . . . [rather than] on a taste for discrimination." Richardson and Gordon (1979) find study areas in predominantly black West Oakland, California not to be "mortgage deficient" relative to surrounding areas.

2. However, it should also be noted that Barth, Cordes, and Yezer (1979) found that the likelihood of a mortgage default was significantly higher if the borrower was black. It is possible that both the Bradbury, Case, and Dunham (1989) results and the Barth, Cordes and Yezer (1979) results are linked to other characteristics, such as income stability, that are correlated with, but not caused by, race. Moreover, in discussing his own work in the field, Canner (1981, p. 68) notes that "[i]t must be emphasized that evidence of racial discrimination in institutional mortgage lending found in this study, as well as others, should be properly viewed outside the redlining process per se. While racial minorities tend to congregate geographically, the redlining process cuts across racial lines."

The post-CRA net return function for the bank, therefore, becomes:

$$\pi = i_P\theta + i_R(1-\theta) - \beta\text{var}[r_P\theta + r_R(1-\theta)] - C_P(I_P)\theta - C_R(I_R)(1-\theta) - d(\bar{\theta}-\theta) - r_d \quad (8)$$

and the optimal portfolio share in neighborhood P loans becomes:

$$\theta_{CRA} = \frac{i_P - i_R + C_R - C_P + 2\beta[\text{var}(r_R)] + d}{2\beta[\text{var}(r_P) + \text{var}(r_R)]} \quad (9)$$

The only difference between (4) and (9) is that d now appears as an argument in the numerator. As the penalty for allocating too little credit to neighborhood P increases, the bank's optimal share of lending to P rises accordingly.

With the imposition of the CRA penalty, where binding, the post-CRA value of θ exceeds the value obtained in Section I by the ratio of the penalty to the weighted sum of the return variances, thus increasing the bank's lending in the poor neighborhood. The bank now treats the CRA penalty as an additional cost of doing business and, in effect, chooses the optimal penalty.

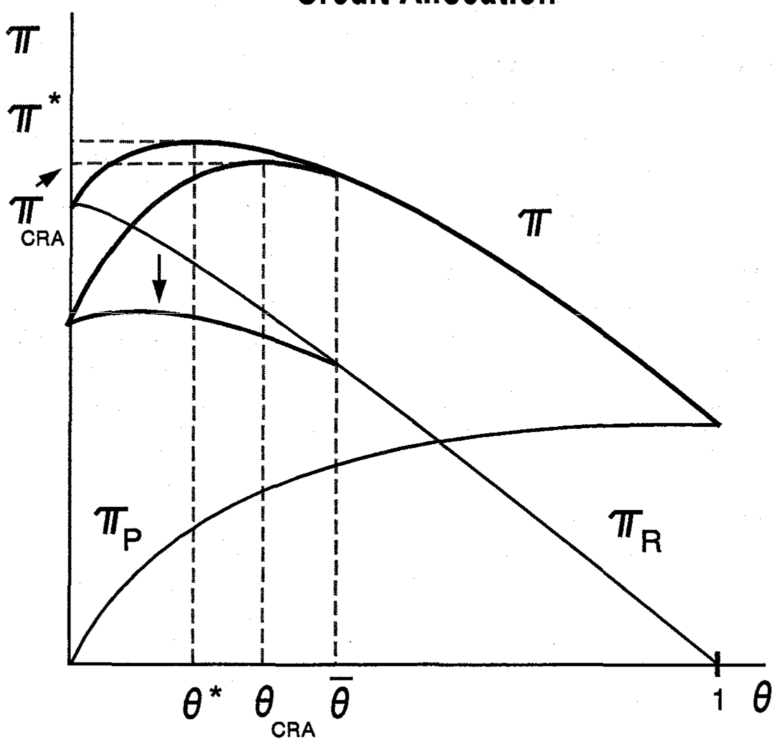
The influence of CRA on the model's solution is depicted graphically in Figure 3. Again, the solid curves represent an initial equilibrium, with θ^* the pre-CRA optimal portfolio allocation. In the presence of the CRA penalty, the returns to neighborhood R loans are reduced for all portfolio allocations below $\bar{\theta}$. The bank must pay the CRA penalty for not allocating a sufficient amount of its portfolio to neighborhood P loans. This is shown in Figure 3 by the downward shift of the π_R function for all values of θ less than $\bar{\theta}$. Our assumption that the penalty increases the further θ^* is from the social optimum is presented by the increasing distance between the black and green curves as θ approaches zero.

As the CRA penalty shifts down a portion of the π_R function, it also shifts the total return function for all values of θ less than $\bar{\theta}$. The profit-maximizing bank chooses the highest point on the total return function, in this case corresponding to the value θ_{CRA}. This portfolio allocation is closer to the socially optimal value of θ than the original solution. CRA thus has the desired impact of raising the proportion of the bank's portfolio allocated to neighborhood P loans. Moreover, the larger the penalty, the larger is the resulting shift in the portfolio allocation.

The achievement of this goal, however, comes at a cost. Total returns for the bank are smaller after the imposition of CRA (as long as the CRA constraint is binding and the actual, as opposed to perceived, risks of lending in neighborhood P are higher than those in neighborhood R). CRA imposes an additional cost on the bank and induces it to increase its lending to the neighborhood with the higher expected return variance. Total returns decline because this

Figure 3

Effect of CRA on Optimal Credit Allocation



higher variance is associated with a greater probability of default. From the bank's standpoint, CRA has an undesirable impact. In essence, CRA forces the bank to pay the social cost of increased lending to the poor neighborhood.[19]

One beneficial side-effect of CRA is its impact on information gathering. Figure 3 does not show the separate information effects that augment the impact of the penalty. Banks seek to reduce the penalty by raising θ, which in turn raises the marginal benefit of information in neighborhood *P*. Banks, therefore, invest more in I_P and less in I_R. The change in information investment shifts the π_P and π_R curves as in Figure 2, reinforcing the increase in θ_{CRA}, and moving it closer to $\bar{\theta}$. Thus, the ability to invest in information moves the equilibrium credit allocation even closer to the socially desired value than is shown in Figure 3.

This informational effect of CRA potentially can mitigate some of the cost of complying with the law. If the bank's initial perception of the low-income neighborhood's risk was too high, CRA's incentive to gather more information can lead the bank to discover that there are far more high-quality loans that can be made in the area than it initially believed. Of course, the increased information also may confirm the bank's initial characterization of the neighborhood's risk. And, in fact, if credit quality is too low, the bank may choose to redline and pay the regulatory penalty.

CRA's effect on information, as modeled here, also achieves the "effort-oriented" enforcement of CRA. Banks are encouraged to aggressively seek loans to boost their portfolio of low-income neighborhood loans, thus satisfying the regulatory interpretation of the law.

IV. Implications and Institutional Developments

One particularly attractive feature of the imperfect information model is its potential to explain the emergence of post-CRA institutional arrangements. The model suggests that socially suboptimal lending occurs in part because the costs of finding good loans are too high, not necessarily because good loans cannot be made in a particular neighborhood.

A predictable response of banks to this new regulatory constraint is to seek ways to minimize the cost and risks of complying with CRA regulations. Several arrangements have emerged in the financial community to reduce the cost of CRA compliance. First, as an explicit response to CRA, financial institutions in a number of states have formed multi-institution consortia[20] which not only lower the per-institution cost of information, but allow participants to share credit risks. A prime example is the California Community Reinvestment Corporation, in which major lenders in the state have pooled funds in a separate entity whose sole directive is to find and make loans in disadvantaged neighborhoods. Since these same individual institutions perform many other types of loan functions in-house, it is clear (and is sometimes explicitly stated) that the establishment of such consortia serves the purposes of lowering per-institution costs of information and of spreading risk in a lending process where such costs and risks are relatively high.

Second, some banks allow non-profit community groups to perform the initial applicant screening for CRA-related loans.[21] This approach lowers the costs of information for the lending institution by shifting part of the search and monitoring costs to the community groups. This approach also may reduce default risks since the community group, by placing its "reputational capital" on the line, has an interest in encouraging the borrower to follow the terms of the loan. It also increases community group sensitivity to the credit-risk problems faced by institutions when lending in certain areas.

Third, some banks form separate corporations for CRA activity which allow banks to take equity positions in the borrower as well as debt positions.[22] Joint loan-equity positions, when they are possible, have been shown to increase the monitoring and information gathering capability of the lender.[23,24]

Two points are particularly interesting with regard to all of these arrangements. First, these strategies are explicitly targeted at CRA and are not widely applied to other lending problems. Second, all of these strategies are aimed at reducing the costs of obtaining neighborhood information and possibly at spreading risk.

The model provides an explanation for these arrangements and an interpretation of their effects on lending to disadvantaged neighborhoods. Consider, first, a post-CRA arrangement in which a number of banks pool their resources to obtain and share information about the poor neighborhood. If we assume that *N* banks participate in the consortium (each as an equal partner), then for each unit of information the individual bank purchases, it receives $N-1$ units through the consortium sharing agreement. In effect, for a given quantity of information, the cost to the individual bank is split among N institutions. The neighborhood P cost function thus becomes $C_P(I_P)/N$. Maximizing the modified adjusted return function yields the following solution for θ:[25]

$$\theta_{POOL} = \frac{i_P - i_R + C_r - \frac{C_P}{N} + 2\beta[\text{var}(r_R)] + d}{2\beta[\text{var}(r_P) + \text{var}(r_R)]} \qquad (10)$$

For a given quantity of I_P, θ is unambiguously larger than in the previous case where each bank obtains neighborhood P information on its own. Clearly, if information costs are lower, the optimal perceived return function for the bank is higher. (In fact, if N becomes so large that per-institution costs become very small, the return function may approach the full-information returns function, and problems caused by imperfect information may be fully mitigated.)

Referring back to the diagrammatic treatment, a decline in information costs acts the same way as that portrayed in Figure 2. Because the marginal cost of a given amount of information falls with pooling, the optimal amount of information gathered rises and the marginal profit function for neighborhood P loans shifts outward, generating an increase in lending to that neighborhood.

Pooling arrangements increase the total amount of information gathered. To see this, consider the first-order condition for neighborhood P information and the resulting partial equilibrium solution for marginal cost:

$$\frac{C'_P(I^*_P)}{N} = -\beta\theta\lambda'_P(I^*_P) \qquad (11)$$

For a given value of neighborhood P information, the marginal cost is lower than in the previous case, and probably lower than the marginal benefit of investing in information. Each bank in the consortium faces a private incentive to invest in more information since they share the costs. The net result is a quantity of I_P that is greater than the case without the consortium. The cost-pooling arrangement thus yields a greater investment in information than the case where all banks operate alone. Such pooled arrangements appear to be particularly cost-effective mechanisms for overcoming problems of imperfect information in lending markets.

Although such methods may minimize the costs of CRA compliance, it is important to emphasize that participating banks still can be expected to be worse off compared to the no-CRA case. If pooling arrangements yielded profits similar to those in the wealthier areas, banks would have had an incentive to form joint ventures before CRA was adopted.

That consortia form under CRA and not without CRA suggests several properties of the cost function. There are economies of scale in information that were not exploited previously. Thus, a bank's costs can be reduced by sharing information, increasing total information while cutting individual information gathering. Moreover, the minimum efficient scale of lending necessary to satisfy CRA may be too high for one bank to enter separately and make normal profits. The consortium, in contrast, may be able to attain a scale of lending activity sufficient to make CRA-related lending profitable. Finally, the costs of forming and maintaining a lending consortium may have prevented pre-CRA arrangements of this type. The benefits of reducing the CRA penalty, however, act to offset these costs.

This result suggests that even with pooling arrangements, profits are lower under CRA. Otherwise, (a) it would not be necessary to form consortia, or (b) if the expected return to lending in the neighborhood is sufficiently high, such consortia would have formed without CRA.

The other arrangements, that is, the involvement of community groups and the use of greater equity control, offer similar advantages in reducing information costs. Community groups may have lower costs of finding good borrowers because of their familiarity with the neighborhood and the potential borrowers. Greater equity control offers lower monitoring costs to the bank, although it does not reduce the initial information cost of finding the loans.

The two most prominent methods, the use of community groups and bank-pooling arrangements differ in their distribution of costs. The use of community groups is attractive to banks since the information costs are passed along to the community and not borne directly by the bank. However, to the extent that banks have a comparative advantage in identifying good loans, that approach may be less efficient. The costs of bank pooling operations are directly borne by the stockholders or customers of the banks, but they have greater control over the lending process.

Other potential arrangements also can be envisioned within this framework. If the costs of complying with social goals put banks at a competitive disadvantage, it might be appropriate to use government funds to subsidize the cost of gathering information, rather than effectively taxing banks as occurs at present. Government agencies could assist community group efforts in screening, or they could subsidize information costs or risks through tax credits. In such ways, the costs of social policy can be directly transferred to society, rather than indirectly through the effect on bank costs, which are borne by bank customers and owners.

V. Summary and Conclusions

In this article, we incorporate CRA policies explicitly in a microeconomic model of bank behavior. We demonstrate that CRA can be viewed as an additional cost or tax on banks when they fail to achieve some socially desirable balance of lending across neighborhoods. Moreover, part of the effectiveness of CRA derives from its inducements to banks to increase lending in disadvantaged neighborhoods—costs that it would not otherwise undertake if CRA were not imposed.

Previous studies have not explicitly modeled the linkage between bank behavior and CRA. Instead, for the most part, they have dealt with the empirical question whether neighborhoods are rationed on the basis of non-economic factors. These studies have been concerned not with *how* CRA affects lending activity, but with *whether* CRA was needed to correct some systematic bias in lending patterns.

In the context of our model, differences in neighborhood lending patterns arise in response to differences in the perceived risks of lending in certain areas. Thus, given the costs involved in obtaining information about neighborhoods, these lending patterns may be optimal from the standpoint of a private, profit-maximizing bank. However, our interpretation of the CRA suggests that it has been invoked to overcome what is viewed as *socially suboptimal* lending in disadvantaged neighborhoods.

The current study provides a better understanding of the relationship between bank behavior and the application of the CRA. Results from this study lead to several conclusions. First, CRA places the social cost of more geographically even lending directly on banks. Banks are, in effect, taxed in such a way as to force them to achieve social goals, with bank customers or shareholders paying the cost of the policy.

Second, by imposing a penalty, CRA will increase lending to disadvantaged neighborhoods—an effect that is reinforced in the model by the assumption of imperfect information. When banks are induced by CRA to increase lending in the poor neighborhood, the value of information about that area rises. The induced investment in information increases the bank's knowledge about the poor neighborhood and may reveal additional low-risk loan projects. Information effects thus raise the proportion of the portfolio allocated to poor neighborhood loans over and above the direct effect of the CRA penalty.

An important policy question arises: if lending patterns are suboptimal from society's standpoint, is it efficient and equitable to place the cost of that social policy on banks? Other mechanisms can be found to lower the cost of information about some neighborhoods, and banks could be subsidized in their information costs rather than taxed.

The policy question raised by this article is whether CRA is the most efficient way to achieve this social goal. Future researchers may wish to shift their focus from whether lending to various neighborhoods is sufficient, to comparing the relative advantages of other mechanisms that can remove the informational inefficiencies that inhibit the desired investing activity.

Moreover, the "penalty function" implicit in CRA offers only one incentive structure for banks to increase their search activity. This approach, which offers vague regulatory penalties on mergers and acquisitions, may not offer the most efficient incentives to banks to increase their lending to disadvantaged neighborhoods. By starting from a clear understanding of the cause of the suboptimal lending—namely, differential risks, exacerbated by costly information—the incentive structure to achieve that goal can be crafted more efficiently.

NOTES

1. See, for example, James (1987).

2. Sullivan and Pozdena (1982) make the distinction between rational redlining, where differential lending occurs as a result of differences in future prospects of borrowers or projects, and irrational redlining. The latter is arbitrary and discriminatory. In our model, only rational redlining occurs.

3. This return includes the explicit interest charged on loans as well as other fees and payments that may be required by the lender. These fees include such items as closing costs, origination fees, prepaid interest, etc. Booth (1990) finds that loan fees play an important role in the pricing of bank commercial loans and that loan fees assume an extensive variety of forms.

4. Although banks have considerable latitude to vary interest rates, collateral requirements, and various fees, some flexibility may be sacrificed to gain advantages of scale through standardization. Moreover, recent work by Jaffee and Russell (1990) suggests that lenders may be limited in their ability to differentiate loan terms because of social pressures regarding "fairness," as well as legal limits associated with discrimination.

5. According to the notion of adverse selection, increases in interest rates may discourage safer borrowers and attract riskier ones who have a lower probability of repaying the loan. The increase in interest rates has an adverse effect on the quality of loan applicants and may actually lead to lower expected returns for lenders (after accounting for defaults). This adverse selection effect limits the extent to which interest rates can reflect loan risk. See the box on Imperfect Information vs. Credit Rationing.

6. In the credit rationing literature, Stiglitz and Weiss (1987) implicitly argue that banks choose to allocate credit sequentially to different borrower classes. The bank chooses a loan rate to each class that maximizes the expected utility (return) from loans to that group, taking into account the adverse selection problem. Consequently, rates differ to different groups, but the differences in rates are not necessarily constructed to yield the same expected return.

7. Although this assumption may appear quite restrictive, relaxing it does not affect the general nature of the results. As long as rates do not fully reflect underlying risks, changes in risk will affect expected returns to the bank. In the model discussed here, changes in variance induced by information do not affect loan rates. This can be generalized easily by making loan rates a direct function of risk, in which case there would be some partially offsetting effect on loan rates. This would reduce the magnitude of the effect on portfolio allocation resulting from a change in risk, but it would not change the direction of the effect.

8. In practice, the volume of loans is an endogenous variable that may be at least partially determined by the bank's deposit-taking activities in the two neighborhoods. For purposes of this discussion, however, the assumption of an exogenous total loan volume is not crucial.

9. Information problems may be more acute in the poor neighborhood if banks have fewer branches in those areas. A smaller deposit base makes branching in poor neighborhoods less profitable. In this way, deposit-taking activity can exacerbate information asymmetries across neighborhoods and impinge on the lending decisions.

10. Although this assumption simplifies the ensuing analysis, it is somewhat restrictive. At the very least, we believe that any covariance between these errors is likely to be positive, a relationship that mitigates somewhat, but does not change, the predictions of the model presented below. Negative cross correlations between neighborhood information sets seem unlikely.

11. This assumption is not unreasonable given that deposit insurance in U.S. banking markets has covered virtually 100 percent of deposits.

12. An alternative reason for the negative relationship between returns and variance has to do with the nature of the debt contract. The probability of default rises with variance, as does the probability of a high payoff to the project owner or borrower. Since lenders cannot receive more than the loan rate when favorable outcomes occur, they are not compensated for the increased probability of default. Thus, lenders' returns have a truncated distribution, the mean of which falls as the variance of returns to project owners rises. This is true even if the rising variance is mean preserving.

13. Clearly, the expected loss parameter β could be an increasing function of the portfolio's variance. In that case, the loss associated with having greater variance would be higher than in the case derived here, further encouraging the movement away from the riskier neighborhood.

14. The share of loans made in neighborhood P also increases as the full-information variance of returns in R rises: as $\sigma_R^2 \to \infty$, $\theta \to 1$. An interesting case arises when the two neighborhoods are identical. In this case, variances of returns in the two areas are the same, as are the information cost functions and the interest rates on loans. In this setting, the bank allocates half of its loan portfolio to each neighborhood. Since there are neither greater loan risks in one neighborhood relative to the other nor greater costs of obtaining information, the bank treats both areas the same in its lending activity.

15. The concavity of these return functions can be shown using the full solution to the model presented in the Appendix. We assume, for simplicity, that the two functions are monotonic. Thus, π_P rises continuously as θ increases while π_R falls as θ rises. It is possible that these two functions might be sufficiently concave to slope downward at their ends, a possibility that only reinforces the results presented here.

16. We assume that the economies of scale in information gathering drop off rapidly, in the sense that the negatively sloped portion of the marginal cost schedule occurs only at very low levels of I_P. In practice, banks would tend to operate on the positively sloped portion of the marginal cost curve, such that an increase in θ increases the amount of information acquired.

17. As long as the marginal benefit of information regarding a particular neighborhood exceeds its marginal cost, additional information will shift the return function from that neighborhood upward. Once information investment reaches the optimum, however, any additional investment in information will reduce returns and the π function will shift downward.

18. These regulators include the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation, the Office of the Comptroller of the Currency, and the Office of Thrift Supervision (formerly the Federal Home Loan Bank Board).

19. Banks may be able to pass some of these costs on to their customers in the form of reduced services, higher fees, etc. If costs are too high, in fact, a bank may choose to drop the neighborhood as part of its market area, thus passing the costs on to the community.

20. See Mannion and Faber (1989) p. 26.

21. See *U.S. News & World Report* (1989), pp. 26-27.

22. See Mannion and Faber (1989), p. 26.

23. See Kim (1989).

24. An alternative strategic response of banks could be to close branches and limit their service area. Some evidence of this effect is presented for the Phoenix metropolitan area by Booth and Smith (1984), where CRA had a negative impact on branching. Limits are imposed on this ability, however, with community groups seeking regulatory prohibitions to such closures.

25. To the extent that pooling also spreads risks and reduces the variance of low-income neighborhood loans to pool participants, $\mathrm{Var}(r_P)$ also may fall in expression (10), further increasing θ.

APPENDIX

The three first-order conditions from the model yield a system of three equations in three variables, θ, I_P, and I_R. We can totally differentiate this 3-equation system and then solve it in terms of the exogenous variables and parameters of the model. The results of this exercise represent the "complete" solutions to the model in that they take account of all interactions among the variables. In matrix notation, the system of three equations can be represented (after total differentiation) as:

$$Ax = y$$

where

$$A = \begin{bmatrix} \{-2\beta[\text{var}(r_R) + \text{var}(r_P)] + d''\} & (-C'_P - 2\beta\theta\lambda'_P) & [C'_R + 2\beta(1-\theta)\lambda'_R] \\ (-\beta\lambda'_P) & (-\beta\theta\lambda''_P - C''_P) & (0) \\ (\beta\lambda'_R) & (0) & [-\beta(1-\theta)\lambda''_R - C''_R] \end{bmatrix}$$

$$x = \begin{bmatrix} d\theta \\ dI_P \\ dI_R \end{bmatrix}$$

$$y = \begin{bmatrix} di_P + di_R - 2\{\text{var}(r_R) - \theta[\text{var}(r_R) + \text{var}(r_P)]\}\, d\beta + 2\beta\theta d\sigma^2_P + 2\beta\,(1-\theta)d\sigma^2_R \\ \theta\lambda'_P d\beta \\ (1-\theta)\lambda'_R d\beta \end{bmatrix}$$

The solution for the vector, x, requires inverting matrix A, i.e.,

$$x = A^{-1}y.$$

The inversion process includes evaluating the determinant of A. In order to ascertain the sign of this determinant, it is necessary to place certain restrictions on the magnitudes of the second order derivatives of the $C(\bullet)$ and $\lambda(\bullet)$ functions in the model. These restrictions entail requiring the second-order own derivatives of these functions (with respect to each neighborhood) to dominate second-order cross effects (i.e., *between* neighborhoods). In addition, we assume that first-order effects generally dominate second-order effects. These restrictions are reasonable and not likely to be violated.

With the above-imposed restrictions, we obtain the following general form:

$$x = A^{-1}y = \begin{bmatrix} (-) & (-) & (+) \\ (-) & (-) & (-) \\ (+) & (-) & (-) \end{bmatrix} [y].$$

This equation system leads to the following set of results:

$$\frac{\partial\theta}{\partial i_P} > 0 \qquad \frac{\partial I_P}{\partial i_P} > 0 \qquad \frac{\partial I_R}{\partial i_P} < 0$$

$$\frac{\partial\theta}{\partial i_R} < 0 \qquad \frac{\partial I_P}{\partial i_R} < 0 \qquad \frac{\partial I_R}{\partial i_R} > 0$$

$$\frac{\partial\theta}{\partial\beta}\ ?\ 0 \qquad \frac{\partial I_P}{\partial\beta} > 0 \qquad \frac{\partial I_R}{\partial\beta} > 0$$

$$\frac{\partial\theta}{\partial\sigma^2_P} < 0 \qquad \frac{\partial I_P}{\partial\sigma^2_P} < 0 \qquad \frac{\partial I_R}{\partial\sigma^2_P} > 0$$

$$\frac{\partial\theta}{\partial\sigma^2_R} > 0 \qquad \frac{\partial I_P}{\partial\sigma^2_R} > 0 \qquad \frac{\partial I_R}{\partial\sigma^2_R} < 0$$

We have also considered the impact of parameters that shift the costs entailed in gathering information (α_i), the underlying return variance (γ_i), and the CRA penalty function (δ).

$$\frac{\partial\theta}{\partial\alpha_P} < 0 \qquad \frac{\partial I_P}{\partial\alpha_P} < 0 \qquad \frac{\partial I_R}{\partial\alpha_P} > 0$$

$$\frac{\partial\theta}{\partial\alpha_R} > 0 \qquad \frac{\partial I_P}{\partial\alpha_R} > 0 \qquad \frac{\partial I_R}{\partial\alpha_R} < 0$$

$$\frac{\partial\theta}{\partial\gamma_P} < 0 \qquad \frac{\partial I_P}{\partial\gamma_P} < 0 \qquad \frac{\partial I_R}{\partial\gamma_P} > 0$$

$$\frac{\partial\theta}{\partial\gamma_R} > 0 \qquad \frac{\partial I_P}{\partial\gamma_R} > 0 \qquad \frac{\partial I_R}{\partial\gamma_R} < 0$$

$$\frac{\partial\theta}{\partial\delta} > 0 \qquad \frac{\partial I_P}{\partial\delta} > 0 \qquad \frac{\partial I_R}{\partial\delta} < 0$$

REFERENCES

Ahlbrandt, Roger S., Jr. "Exploratory Research on the Redlining Phenomenon," *American Real Estate and Urban Economics Association Journal* 5 (Winter 1977), pp. 473-81.

Avery, Robert B. and Thomas M. Buynack. "Mortgage Redlining: Some New Evidence," Federal Reserve Bank of Cleveland *Economic Review*, Summer 1981.

Barth, James R., Joseph J. Cordes, and Anthony M. J. Yezer. "Financial Institution Regulations, Redlining, and Mortgage Markets," in *The Regulation of Financial Institutions*, Conference Series 21, Federal Reserve Bank of Boston (1979), pp. 101-43.

Benston, George J. "Mortgage Redlining Research: A Review and Critical Analysis," *Journal of Bank Research*, Spring 1981, pp. 8-23.

Booth, James. "The Determinants of Spreads on Bank Loan Contracts," unpublished manuscript, 1990.

Booth, James R. and Richard L. Smith, II. "The Impact of the Community Reinvestment Act on Branching Activity of Financial Institutions," *Journal of Bank Research* 15 (Summer 1984), pp. 123-28.

Bradbury, Katharine L., Karl E. Case, and Constance R. Dunham. "Geographic Patterns of Mortgage Lending in Boston, 1982-1987," *New England Economic Review*, September/October 1989, pp. 3-30.

Calem, Paul S. "The Community Reinvestment Act: Increased Attention and a New Policy Statement," Federal Reserve Bank of Philadelphia *Business Review*, July/August 1989, pp. 3-16.

Canner, Glenn B. "Redlining and Mortgage Lending Patterns," in J. Vernon Henderson, ed., *Research in Urban Economics: A Research Annual*. Greenwich, Connecticut: JAI Press (1981), pp. 67-101.

_________. "The Community Reinvestment Act: A Second Progress Report," *Federal Reserve Bulletin* 67 (November 1981), pp. 813-823.

_________. *The Community Reinvestment Act and Credit Allocation*. Washington, D.C.: Board of Governors of the Federal Reserve System Studies 117 (1982).

_________. *Redlining: Research and Federal Legislative Response*. Washington, D.C.: Board of Governors of the Federal Reserve System Studies 121 (1982).

_________, and Joe M. Cleaver. "The Community Reinvestment Act: A Progress Report," *Federal Reserve Bulletin* 66 (February 1980).

Dingemans, Dennis. "Redlining and Mortgage Lending in Sacramento," *Annals of the Association of American Geographers* 69 (June 1979), pp. 225-39.

Hutchison, Peter M., James R. Ostas, and J. David Reed. "A Survey and Comparison of Redlining Influences in Urban Mortgage Lending Markets," *American Real Estate and Urban Economics Association Journal* 5 (Winter 1977), pp. 463-72.

Jaffee, Dwight and T. Russell. "Fairness and the Structure of the Loan Market," unpublished draft manuscript, 1990.

Jaffee, Dwight and Joseph Stiglitz. "Credit Rationing," in B. M. Friedman and F. H. Hahn, eds., *Handbook of Monetary Economics*, Volume II. New York: Elsevier Science Publishers, 1990, pp. 837-88.

James, Christopher. "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* 19 (1987), pp. 217-235.

Jones, Colin and Duncan MacLennan. "Building Societies and Credit Rationing: An Empirical Examination of Redlining," *Urban Studies* 24 (June 1987), pp. 205-216.

Kim, Sun Bae. "Modus Operandi of Lender-Cum-Shareholder Banks," Department of Economics, University of Toronto, unpublished paper, November 1989.

Mannion, Robert E. and Michael Faber. "Reinvigorated CRA Requires Greater Attention," *ABA Bank Compliance* 44 (Autumn 1989), pp. 21-26.

National Peoples Action. *Survey of Redlining in Major Cities*. Testimony submitted in hearings before U.S. Senate. November 23, 1976.

New York Public Interest Research Group. *Take the Money and Run: Redlining in Brooklyn*. Testimony submitted in hearings before U.S. Senate. March 25, 1977.

Richardson, Harry W. and Peter Gordon. "Measuring Mortgage Deficiency and Its Determinants," *The Annals of Regional Science* 13:3 (November 1979), pp. 25-34.

Stiglitz, Joseph E. and Andrew Weiss. "Credit Rationing in Markets with Imperfect Information," *American Economic Review* 73(3) (June 1981), pp. 393-410.

_________. "Credit Rationing With Many Borrowers," *American Economic Review* 77(1) (March 1987), pp. 228-231.

Sullivan, Alane K. and Randall J. Pozdena. "Enforcing Anti-Redlining Policy Under the Community Reinvestment Act," Federal Reserve Bank of San Francisco *Economic Review*, Spring 1982, pp. 19-34.

U.S. News & World Report. "A Housing Program That Really Works." February 27, 1989, pp. 26-27.